MST121 Chapter D1

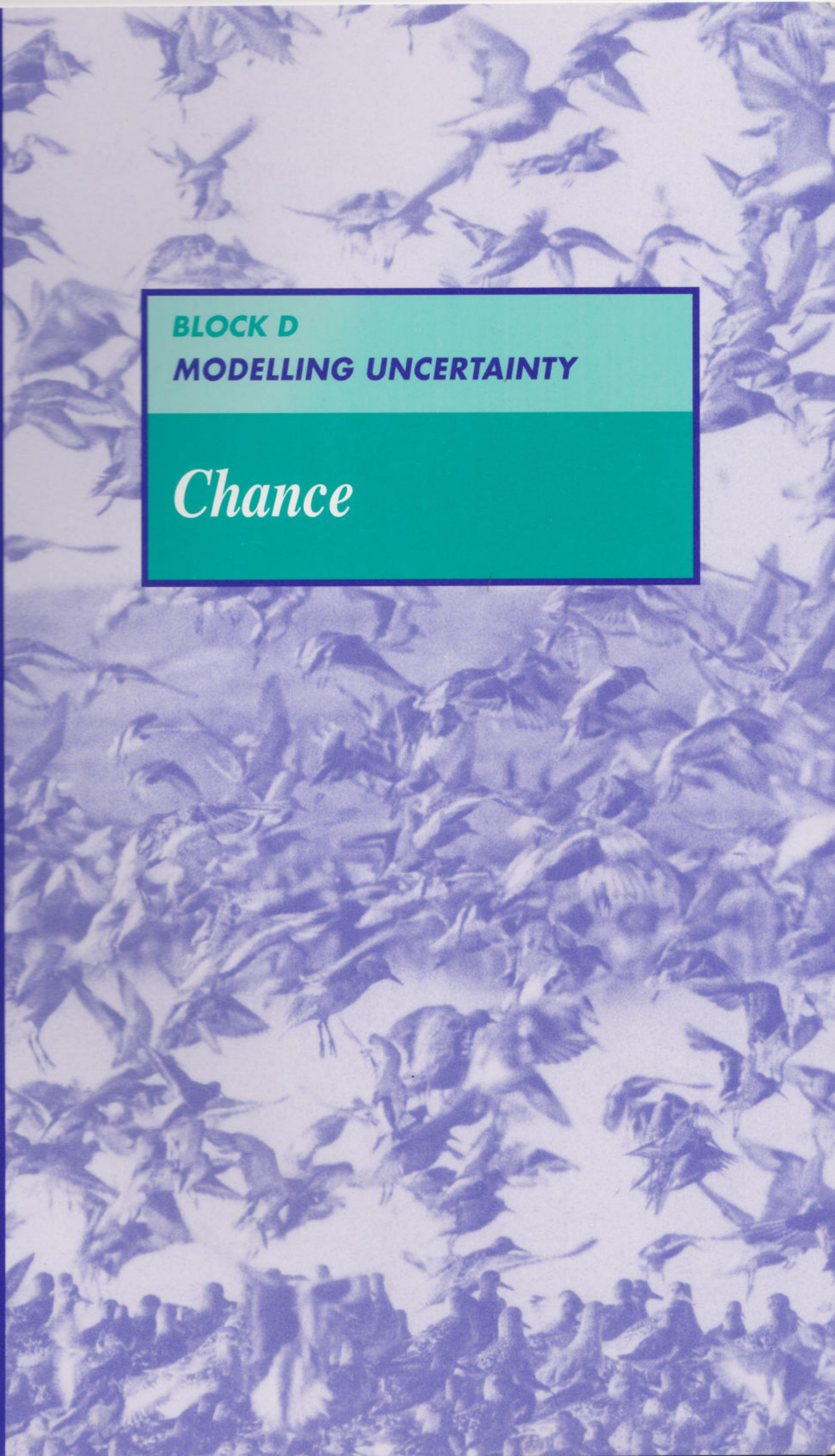
The Open University

A first level interdisciplinary course

Using Mathematics

CHAPTER

D1



BLOCK D

MODELLING UNCERTAINTY

Chance

MST121 Chapter D1

The Open University

A first level interdisciplinary course

Using Mathematics

BLOCK D
MODELLING UNCERTAINTY

Chance

Prepared by the course team

CHAPTER D1

About this course

This course, MST121 *Using Mathematics*, and the courses MU120 *Open Mathematics* and MS221 *Exploring Mathematics* provide a flexible means of entry to university-level mathematics. Further details may be obtained from the address below.

MST121 uses the software program Mathcad (MathSoft, Inc.) and other software to investigate mathematical and statistical concepts and as a tool in problem solving. This software is provided as part of the course.

This publication forms part of an Open University course. Details of this and other Open University courses can be obtained from the Course Information and Advice Centre, PO Box 724, The Open University, Milton Keynes, MK7 6ZS, United Kingdom: tel. +44 (0)1908 653231, e-mail general-enquiries@open.ac.uk

Alternatively, you may visit the Open University website at http://www.open.ac.uk where you can learn more about the wide range of courses and packs offered at all levels by The Open University.

To purchase a selection of Open University course materials, visit the webshop at www.ouw.co.uk, or contact Open University Worldwide, Michael Young Building, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom, for a brochure: tel. +44 (0)1908 858785, fax +44 (0)1908 858787, e-mail ouwenq@open.ac.uk

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 1997. Second edition 2004. Reprinted 2005.

Edited, designed and typeset by The Open University, using the Open University TEX System.

Printed in the United Kingdom by Thanet Press Ltd, Margate.

ISBN 0 7492 6661 9

2.2

Contents

Introduction to Block D

> Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.
>
> H. G. Wells (1866–1946)

For centuries, the general scientific and philosophical view was that all the variability in a phenomenon could be explained if only all the factors which cause the variation could be identified: it was believed that, once all relevant factors had been identified, a model could be developed to explain the variation and to predict what will happen. However, in many situations, to identify all the variables involved and to develop a model which includes them all can be a complicated if not impossible task.

By the 18th century, the need for models which incorporate uncertainty was recognised. For instance, Gauss (1777–1855) and Laplace (1749–1827) both recognised the usefulness of using the notion of chance to model imprecision in measurements. Since that time, models incorporating uncertainty have been applied in many different fields – insurance, nuclear physics, genetics and astronomy, to name a few. Models involving chance have also been developed for the spread of an epidemic and the behaviour of a queue.

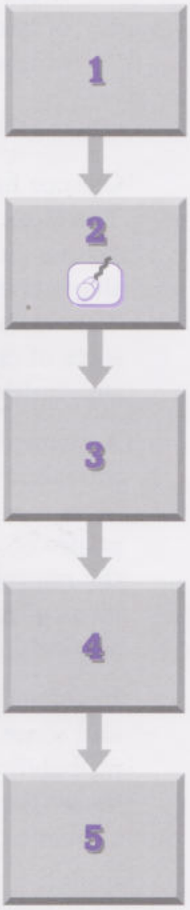
The title of this block is *Modelling uncertainty*. In Chapter D1, we begin by considering a concept which is fundamental to all models for chance events and which underpins statistical thinking: *probability*. In Chapters D1 and D2, two models for the variation observed in a variable are discussed – one model for a discrete variable and one for a continuous variable. A key factor in the development of statistical thinking was the desire to use information gained from a sample to make inferences about the population from which the sample was taken. Chapters D3 and D4 are concerned with drawing inferences about populations from samples of data; these chapters look at three types of statistical investigation. Chapter D3 looks at estimating an unknown quantity; the first part of Chapter D4 investigates differences between populations by comparing samples of data; and the second part is about looking for relationships between variables.

The first step in any statistical investigation is to specify its purpose and pose a precise question. Once this has been done, relevant data are collected. You will recognise these as two aspects of the first stage of the modelling process: specifying the purpose. The next step is to analyse the data – this is the 'doing the mathematics' stage of the modelling process – and then the results are interpreted. Chapters D3 and D4 concentrate on posing a precise question, analysing the data and interpreting the results; you will not be asked to collect the data yourself – the data are provided.

When analysing data that have been collected, statisticians tend to use software which has been designed for this purpose. Mathcad is not designed as a statistical analysis tool, so in this block, instead of using Mathcad you will be using the statistics software package which has been supplied to you with the materials for this block. ~~One module of this software package, called StatsAid, is a computer-mediated learning package designed to support Block D. It explains most of the statistical prerequisites of the course, covering some of the statistical topics taught in the course MU120. See the Guidance notes for Computer Book D for further information.~~

Study guide

This chapter contains five sections, which are intended to be studied consecutively in four study sessions, and an appendix. The first four sections should take one and a half to two and a half hours of study each. Section 5 is relatively short. You will need access to your computer and Computer Book D for Section 2, which contains only computer-based work.



The pattern of study for each session might be as follows:

Study session 1: Section 1.

Study session 2: Section 2.

Study session 3: Section 3.

Study session 4: Sections 4 and 5.

Sections 4 and 5 could be split into two study sessions.

Before studying this chapter, you should be familiar with the following topics, ~~which are covered in the software package StatsAid~~:

◇ the mean of a batch of data;

◇ frequency diagrams.

The mean is ~~also~~ covered in the *Revision Pack*; both topics are covered in the course MU120.

Introduction

> It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge.
>
> *Théorie analytique des probabilités*
> Pierre Laplace (1749–1827)

Chance has been an accepted part of life for many thousands of years. The fears and uncertainties associated with people's experiences of birth, sickness and death, wars, earthquakes, droughts and floods have helped to shape their sense of the unpredictable nature of these life events. Throughout recorded history, people have sought explanations for these sorts of chance events, and linked them to a variety of beliefs and superstitions, many of which continue to flourish today.

A concept which is essential for modelling chance events is that of probability: a probability is basically a number which measures the chance of an event occurring. The desire of some gamblers to analyse various games of chance, particularly ones involving dice, provided the stimulus for the efforts which eventually led to the development of the fundamental ideas of probability theory. Once developed, these ideas were very rapidly applied in many other fields.

In this chapter, some fundamental ideas about probability are introduced, and some of the original problems which prompted the development of probability theory are described and analysed. You will see how these ideas can be applied to model some other situations involving an element of chance: for example, the births of boys and girls.

Section 1 begins with some early history of games of chance, and then the probability of an event is defined. Before any ways of calculating probabilities are introduced, you will be asked to consider a number of questions and to record your ideas about the chances of various events. In the computer section which follows, you will be invited to explore some of these questions and to compare the results you obtain with the ideas you noted when using only your intuition. You will also be asked to make hypotheses about a number of the questions on the basis of your explorations using the computer. In the remaining sections, some basic rules of probability will be introduced and then applied to answer the questions raised in Section 1. You will be able to compare your intuitions and hypotheses with the results obtained using probability theory.

1 Questions of chance

How did the development of a theory for quantifying chance come about? How is chance measured? Are your intuitions about chance events reliable? These are the questions which underlie the material in this section.

In Subsection 1.1, some early history of games of chance is discussed briefly. A definition of the *probability* of an event – a number which quantifies how likely an event is to occur – is given in Subsection 1.2. And in Subsection 1.3 you are asked to use your intuition and experience to propose answers to a number of problems involving chance.

1.1 Games of chance – some history

Games of chance have been around for a very long time. Boards and counters dating back to 3500 BC have been found in Egypt. There are tomb-paintings which suggest that some games from that era involved moving counters on the throw of an *astragalus*. (An astragalus is a bone from the heel of an animal.) The painting in Figure 1.1 shows a nobleman in the after-life using an astragalus in a board game.

Photograph reproduced courtesy of the Oriental Institute of the University of Chicago.

Figure 1.1 Ancient Egyptian wall painting – tomb of Neferronpe

The shape of an astragalus is such that when it is thrown, it can land in one of four positions (it has four fairly flat sides). The astragalus was almost certainly the forerunner of the six-sided die of later times which eventually replaced it. Astragali and dice were both in common use for many centuries. Early dice were roughly hewn and uneven in shape, and no two astragali were the same. So experience gained using one astragalus or die could not be used to predict reliably how another might behave. Thus, for a long time there was no impetus for developing a theory to explain the nature of such chance events as the throw of an astragalus or

the roll of a die. A sketch of an astragalus from a sheep is shown in Figure 1.2, and a six-sided die in Figure 1.3.

Figure 1.2 Astragalus from a sheep

Figure 1.3 Six-sided die

It is not possible to say when gambling originated, but gaming, that is, gambling on the outcomes of games of chance, was widespread in the Roman Empire; it was the common recreation of the time among all sections of society. In the centuries which followed the fall of the Roman Empire, gambling and games of chance continued to flourish in Europe, in spite of the vigorous opposition of the Christian Church. By this time, dice were well made and so some of the most inveterate gamblers must have had some intuitive idea about the relative chances of the different outcomes in the dice games they played. However, throughout this period, the Church regarded secular learning with deep suspicion, and it was not until much later that serious attempts were made to understand and quantify chance.

When a die is rolled, there are six possible outcomes – the scores $1, 2, \ldots, 6$. There is evidence that an understanding of the concept of equally-likely outcomes on the roll of a die had been achieved by the 15th century, and from that time on, efforts were made to explain differences that were observed in the relative frequencies of the various outcomes in dice games. In the 16th century, Girolamo Cardano, a scholar and gambler, made the step from observation to theory. He wrote the following about the outcomes of the roll of a die in his book *Liber de Ludo Aleae.*

The *Liber de Ludo Aleae* (Book on games of chance) was found among Cardano's papers after his death, but was not published until 1663, about a hundred years after it was written.

> One-half the total number of faces always represents equality; thus the chances are equal that ... one of three points will turn up in one throw. For example, I can as easily throw one, three or five as two, four or six. The wagers therefore are laid in accordance with this equality.

In later chapters of the book, he discussed the results of rolling two and three dice, and went on to calculate the odds of various outcomes. He also made calculations for a number of card games.

If one event occurs, on average, four times for every three occasions on which a second event occurs, then we say that the *odds* are four to three in favour of the first event.

The beginnings of modern probability theory are often attributed to the two French mathematicians Blaise Pascal and Pierre de Fermat. Between 1654 and 1660, they corresponded about a number of mathematical problems, including several concerning analysing odds in games of chance. This correspondence seems to have started when the Chevalier de Méré, a French nobleman and enthusiastic gambler, consulted Pascal about a problem that had arisen in a game of chance. Essentially, the question was about how the stakes should be divided between two players when their game is interrupted before either player has obtained the number of points required to win; this problem is often referred to as the Problem of Points. Solutions to this and another problem posed by de Méré can be found in the letters between Pascal and Fermat which survive to this day.

These letters also contain discussion of a number of other mathematical problems.

All the early efforts to solve problems about games of chance were complicated by the absence of the idea of using a probability to measure the chance of an event occurring. The work that had been done on such problems was fragmentary, and there was no established method for

tackling them. The arguments used in solutions were often lengthy, and frequently involved much use of ratio and proportion in order to calculate odds. The idea of a probability seems to have emerged in the latter part of the 17th century. By the time that James Bernoulli wrote his treatise *Ars Conjectandi* (The Conjectural Art) in the 1680s and 1690s, it appears to have become an accepted idea. With the advent of probability, many problems that had been complicated to solve using odds became relatively straightforward, and the explosion in the development and application of a more general theory of chance dates from this time.

This work was not published until 1713, eight years after Bernoulli's death.

1.2 What is probability?

The essence of a chance event is that you do not know whether or not it will happen. Nevertheless, in one particular sense, chance events can be regarded as predictable! Suppose, for instance, that a coin is to be tossed a large number of times. Although you cannot say whether it will land heads up or tails up on any particular toss (you can only guess), you can predict fairly accurately the *proportion* of times that it will land heads up. Since there is no reason to believe that either of heads or tails is more likely to occur than the other, you would expect the coin to land heads up approximately half of the time.

Table 1.1 shows the results of the first 8 tosses in a sequence of 30 tosses of a pound coin. The second row of the table shows the outcome of each toss – h for heads and t for tails. The third row shows the total number of heads obtained so far; and the fourth row shows the proportion of heads so far, that is, the total number of heads so far divided by the number of tosses so far. The final row gives these proportions as decimals.

Table 1.1

Toss number	1	2	3	4	5	6	7	8
Outcome (h or t)	h	t	h	t	t	t	t	h
Number of heads so far	1	1	2	2	2	2	2	3
Proportion of heads (P)	1	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{2}{4}$	$\frac{2}{5}$	$\frac{2}{6}$	$\frac{2}{7}$	$\frac{3}{8}$
P as a decimal	1	0.5	0.667	0.5	0.4	0.333	0.286	0.375

Figure 1.4 shows a plot of P, the proportion of heads so far, on the vertical axis, against the toss number on the horizontal axis. Successive points have been joined with straight lines to show more clearly how the proportion of heads changes as the number of tosses increases.

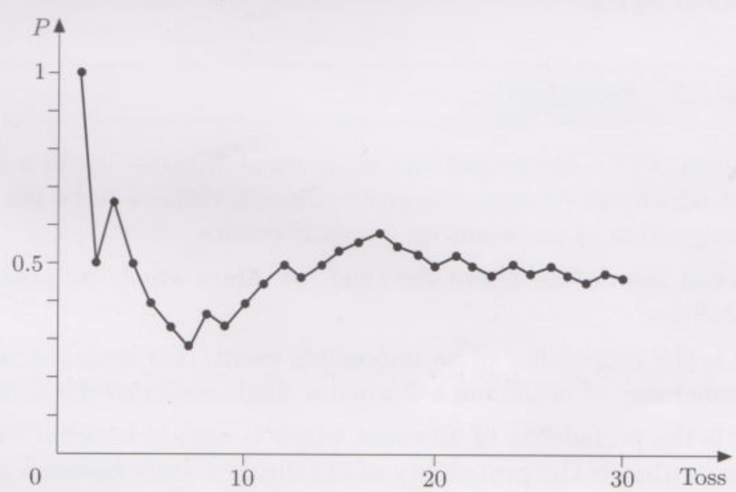


Figure 1.4 Proportion P for 30 tosses of a coin

Activity 1.1 Heads and tails

Take a coin and toss it 30 times, keeping a record of the result of each toss in a table similar to Table 1.1. (Your results will almost certainly be different from those shown in Table 1.1 and Figure 1.4.)

Now plot your results on a graph similar to that in Figure 1.4.

What do you notice about the way the proportion of heads changes as the number of tosses increases?

Comment

An interesting phenomenon is apparent from Figure 1.4, which you may also have observed for your results. For small numbers of tosses, there are quite large fluctuations in the proportion of heads observed. However, as the number of tosses increases, the differences between successive values of P tend to become smaller: the proportion of heads observed seems to be settling down to some constant value. The value towards which the proportion of heads is tending is called the **probability** of obtaining a head when the coin is tossed. From Figure 1.4, it looks as though this value may be $\frac{1}{2}$. Of course, it is possible that, in a sequence of only 30 tosses, the proportion of heads differs substantially from $\frac{1}{2}$: you may have obtained a value either greater than or less than $\frac{1}{2}$. To be confident that the proportion of heads really does approach the value $\frac{1}{2}$, a much longer sequence of tosses is required. However, tossing a coin a large number of times would be a lengthy and tedious exercise: in the computer section, you will be able to use the computer to *simulate* tossing a coin a large number of times, to see what might happen if you actually carried out the tossing.

The idea of assigning a number to an event which expresses how likely that event is to occur is fundamental to probability theory. In general, suppose that an event E (say) may or may not occur in an experiment, and that the experiment can be repeated as often as we like. For instance, the event E might be obtaining a head when a coin is tossed, or a six when a die is rolled. If the experiment is repeated many times, then the observed proportion of occasions on which the event E occurs will tend to settle down to some constant value as the number of times the experiment is repeated increases. This value is called the **probability** of the event E and is denoted $P(E)$.

$P(E)$ is usually read as 'the probability of E' or simply as 'P of E'.

Activity 1.2 Probabilities

As just stated, $P(E)$, the probability of an event E occurring in a single experiment which can be repeated many times, is defined to be the long-run proportion of occasions on which E occurs.

(a) What can you deduce about the range of values which are possible for probabilities?

(b) What is the probability of an impossible event? For example, what is the probability of obtaining a 7 when a single six-sided die is rolled?

(c) What is the probability of an event which is certain to occur? For example, what is the probability of obtaining a score between 1 and 6 inclusive when a single six-sided die is rolled?

Comment

(a) Since a probability is a proportion (the proportion of occasions on which an event occurs), it is always a number between 0 and 1.

(b) An impossible event never happens (you never score 7 with a six-sided die), so its probability – the proportion of occasions on which it occurs – is 0.

(c) Similarly, an event which is certain to happen always occurs (you always get a score between 1 and 6 when you roll a six-sided die), so its probability is 1.

We can summarise the results of Activity 1.2 as follows.

> **1.** For any event E,
>
> $$0 \leq P(E) \leq 1.$$
>
> **2.** If an event E never happens, then $P(E) = 0$.
>
> **3.** If an event E is certain to happen, then $P(E) = 1$.

The first property is a very useful one to remember. You can use it as a 'commonsense' check in probability calculations: if a calculation results in a 'probability' outside the range 0 to 1, then you know you have made a mistake.

We have now defined what is meant by the probability of an event. But how can we calculate probabilities in practice? In general, it is not feasible to carry out repeated experiments to estimate probabilities; and sometimes it is impossible – for instance, how could you calculate the probability that you will be involved in a motor accident within the next year? However, for coin-tossing, if we believe that the two possible outcomes, heads and tails, are equally likely (and nothing else is possible), then we can predict that the proportion of tosses resulting in heads will be approximately $\frac{1}{2}$. Because of the symmetry of a coin, we are able to say, without carrying out a long sequence of tosses, that the probability of a head is $\frac{1}{2}$. Using the notation just introduced, this is written as $P(h) = \frac{1}{2}$.

The idea of equally-likely outcomes can be used to calculate probabilities in many other situations. It is fundamental to many of the examples discussed in later sections of this chapter. Problems involving dice, for instance, can be tackled by assuming that, when a die is rolled, each of the six faces is equally likely to be uppermost when it lands. In a large number of rolls of a single die, we would expect each face to be uppermost for approximately $\frac{1}{6}$ of the rolls, that is, approximately $\frac{1}{6}$ of the time: so

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \tfrac{1}{6}.$$

And, since 3 out of the 6 equally-likely outcomes are even numbers, we would expect an even number $\frac{3}{6}$ of the time, so

$$P(\text{even number}) = \tfrac{3}{6} = \tfrac{1}{2}.$$

The next two activities relate to situations where it may be assumed that all the possible outcomes are equally likely.

Activity 1.3 Deck of cards

A standard pack of 52 playing cards consists of four suits – hearts, clubs, diamonds and spades. Each suit contains thirteen cards – an ace, cards numbered 2 to 10, a jack, a queen and a king. A pack of 52 cards is shuffled thoroughly and the top card is turned face up. Write down the probability that this card is

(a) the ace of spades,

(b) an ace,

(c) a heart.

A solution is given on page 55.

Activity 1.4 Lotteries

(a) In the 1970s, the state of New Jersey in the USA had a lottery with a single 50 000 dollar prize. One million tickets were sold: the tickets were numbered from 000 000 to 999 999. The winning ticket was identified by choosing a six-digit number at random, that is, in such a way that each six-digit number had an equal chance of being selected.

What is the probability that a person will win such a lottery if he or she buys (i) one ticket, (ii) ten tickets?

(b) In the British National Lottery, which was introduced in 1994, a player chooses six different numbers between 1 and 49 (inclusive). Each week six numbers are drawn at random, and a player wins a share of the jackpot if all his or her six numbers match the six numbers drawn. There are 13 983 816 different selections of six numbers between 1 and 49 and, since the six numbers are drawn at random, the selections are all equally likely to occur.

For MST121, you are not expected to know how to calculate the number of different selections; this is calculated in MS221 Block B.

Find the probability that a player will win a share in the jackpot if he or she makes (i) one selection, (ii) ten different selections, (iii) 100 different selections.

A solution is given on page 55.

Unfortunately, it is not always possible to calculate probabilities simply by considering equally-likely outcomes. For instance, you could not use this method to find the probability that when a drawing pin is dropped it will land point up; or to find the probability that it will snow this year in London on Christmas Day; or to find the probability that a person taking out health insurance will make a claim in the next year. In such cases, we must return to the definition of a probability as the long-run proportion of the time that an event occurs. In the case of the drawing pin, we could toss it a large number of times and hence estimate the probability that in a single toss it will land point up.

Although we cannot carry out a large sequence of experiments to estimate the other two probabilities, we can make use of existing data. For instance, we could find out how many times snow has fallen in London on Christmas Day in the last hundred years, and use the proportion of years on which snow has fallen as an estimate of the probability that it will snow in London on Christmas Day this year. Similarly, an insurance company could estimate the chances of a potential customer making a claim using models based on information about, amongst other things, the claims records of similar policy holders; a decision could then be made about whether to issue a policy and what premium to charge.

The idea of estimating probabilities from data has been accepted for several hundred years. In the 17th century, the Englishman John Graunt used data from the weekly Bills of Mortality to calculate empirical probabilities of various life events – for example, of dying from a particular disease, or in an accident, or in childbirth.

He also estimated the population of London at risk in a number of plague years, and compared the severity of the different epidemics by estimating the proportion of the population who died of the plague in each year. He concluded that 1603 was the worst plague year (about half of the population at risk died), and that this outbreak was much more severe than the 'Great Plague of London' of 1665 (in which about a quarter of the population at risk died).

Modern statistical theory and practice has its roots in the two approaches to probability which have been discussed briefly in this subsection – the theoretical approach based on equally-likely outcomes, and the empirical approach based on the collection of data. The empirical approach was developed in England during the same period that European mathematicians, as a result of efforts to analyse games of chance, were developing a theory of probability.

Empirical and theoretical probability

In this subsection, you have been introduced to two approaches to probability: empirical and theoretical. A dictionary definition of 'empirical' might be something like the following:

> ... (of knowledge) based on observation or experiment, not on theory.

So an empirical approach might involve taking observations from the past – whether or not it snowed in London on Christmas Day, say – and using the proportion of days that it snowed as an estimate of the probability that it will snow on Christmas Day this year. Or it might involve repeating an experiment many times – say, tossing a coin or rolling a die 100 times – and noting the outcomes.

Using an empirical approach, if you rolled a die a large number of times and got a six roughly $\frac{1}{6}$ of the time, then you would conclude that the probability of getting a six is approximately $\frac{1}{6}$.

EMPIRICAL APPROACH

THEORETICAL APPROACH

A theoretical approach is based on the geometric symmetry of the coin or die. A coin has two faces, so, assuming that each face is equally likely to occur, each of the two outcomes (head and tail) has a probability of $\frac{1}{2}$. Using a similar argument, a die has six 'identical' faces, so, assuming that the six outcomes (1 to 6) are equally likely to occur, each outcome has a probability of $\frac{1}{6}$.

Theoretical probabilities are calculated by making assumptions (such as those above based on the symmetry of a coin or die), while empirical probabilities are based on experimental evidence or on records of what has happened in the past.

Activity 1.5 Reading probabilities

In this subsection, the basic notation for the probability that an event occurs has been introduced: $P(E)$ is the probability that an event E occurs, and is sometimes said simply as 'P of E'. There have already been quite a number of statements in the text which have used this notation, including the following:

$$P(E) = 0, \quad P(E) = 1, \quad 0 \leq P(E) \leq 1,$$
$$P(h) = \tfrac{1}{2}, \quad P(3) = \tfrac{1}{6}, \quad P(\text{even number}) = \tfrac{3}{6}.$$

What words do you say to yourself as you read these statements? For example, how do you read the first statement? While the notation is new to you, you will probably find it helpful to use words that convey the full meaning of the statements; for example, you might read the first statement as 'the probability that the event E occurs is zero'. However, later on, when you are more familiar with the language and ideas of probability, you may find that it is enough for you to say 'P of E is zero'.

1.3 The questions

This subsection consists of a selection of problems and puzzles. You are not expected to be able to solve the problems at this stage, but wherever possible use your experience and intuition to 'guess' an answer. Write down your 'guesses', so that later you can compare them with the answers to the problems. You will be able to explore some of the problems in the computer section which follows. The tools for tackling the problems are developed in the remaining sections of this chapter, and each problem will be revisited as the necessary ideas and techniques are introduced. You may find the answers to some of the questions surprising, so do not worry too much about whether your intuitive responses are correct. Finding out which of your intuitions are in need of possible revision, and why, will help you to gain a better understanding of the nature of chance events.

The Brains Trust

Some years ago, BBC Radio broadcast a regular programme called the *Brains Trust*. Each week, a selection of questions which had been sent in by listeners were put to a panel of 'brains'. In one programme during the Second World War, the panel was asked: 'What is the law of averages?' One member of the panel, Dr C. E. M. Joad, replied: 'The law of averages says that if you spin a coin a hundred times, it will come down heads fifty times, and tails fifty times.'

Do you think this is correct? If in doubt, try tossing a coin a hundred times to see what happens! And if you do get fifty heads and fifty tails, try repeating the experiment! What do you understand by the phrase 'the law of averages'? How would you use coin-tossing to explain 'the law of averages'?

D'Alembert's heads

The Frenchman Jean d'Alembert was one of the great mathematicians of the 18th century. In 1754, the following problem was proposed to him: in two tosses of a coin, what is the probability that the coin will land heads at least once? He argued that there are three cases: heads on the first toss, heads on the second toss, and heads on neither toss. Two of these three give at least one head; therefore, he argued, the probability required is $\frac{2}{3}$.

What do you think? Do you agree with d'Alembert's argument?

The three-card game

An entertainer at a fairground invites members of the public to bet 50p on the outcome of a three-card game. Three cards are put into a hat: one is white on both sides, one is red on both sides, and the third is white on one side and red on the other. If you decide to play the game, then he lets you choose a card from the hat without looking at it, and place it flat on a table. If the side showing is red, then he says: 'This isn't the white-white card, so it must be one of the other two. I'll bet you 50p that the other side is red.' (So, in this case, he will pay you 50p if the other side is white.) And similarly, if the side showing is white, he offers to bet you 50p that the other side is white.

Would you accept the wager? Do you think it is a fair bet?

Galileo and the three-dice problem

Galileo Galilei was born in Pisa in 1564. He was educated in a Jesuit monastery until he was sixteen, and then spent a short time in commerce. After this, he studied medicine at the University of Pisa and, at the age of 25, became a professor of mathematics there. He later moved to Padua and, in 1613, from there to Firenze (Florence). At this time of his life, Galileo was almost entirely occupied with astronomy, but at some time between 1613 and 1623, he seems to have been instructed to look at a problem concerning the total score obtained when three dice are rolled. In his own words, he was 'ordered to produce' whatever occurred to him about the problem. (Presumably he was asked to look at the problem by his employer/patron, the Grand Duke of Tuscany.)

GALILEO WAS ALSO KNOWN AS THE 'LEANING TOWER OF PIZZA'

The problem was essentially as follows. There are six 3-partitions of 9; that is, there are six different sets of three die scores which add up to 9, namely (621), (531), (522), (441), (432), (333). There are also six 3-partitions of 10: (631), (622), (541), (532), (442), (433). Galileo was asked to investigate why, even though there are the same number of 3-partitions of 9 as there are of 10, 10 seems to be 'more advantageous' in practice. (Presumably the Grand Duke regarded a total of 10 as more advantageous because he had observed that it occurred more often than a total of 9.)

What do you think about this problem? Is a total of 10 more likely than a total of 9 when three dice are rolled? Write down your ideas. We shall return to this problem in Section 3, where you will see how Galileo tackled it.

The Chevalier de Méré: sixes and double-sixes

In his correspondence with Pierre de Fermat, Blaise Pascal raised a problem which had been brought to him by the Chevalier de Méré. According to Pascal, the Chevalier claimed to have 'found falsehood in the theory of numbers'. The Chevalier made this claim because two wagers, which he had reasoned to be equally advantageous, had proved not to be so in practice. He had correctly calculated the odds of rolling at least one six with four rolls of a single die, and found that they are favourable (that is, the probability of doing so is greater than $\frac{1}{2}$). According to Pascal, the Chevalier reasoned that, since '24 is to 36 (which is the number of pairings of the faces of two dice) as 4 is to 6 (which is the number of faces of one die)', it should also be advantageous to bet on rolling a double-six in twenty-four rolls of two dice.

Which do you think is the more likely: rolling at least one six with four rolls of a single die, or rolling at least one double-six with twenty-four rolls of a pair of dice? Or do you agree with the Chevalier that they should be equally-likely events? Could it be that he simply had a run of bad luck at the gaming tables? Or was he right to suspect that the second event was less likely than the first?

Balanced families

The Watsons regard one boy and one girl as the ideal family. When they married, they reasoned that since boys and girls are equally likely, they had an even chance of getting one boy and one girl in their planned family of two.

For the Johnsons, two boys and two girls is the ideal family. They also reckoned that, because boys and girls are equally likely, their chances of achieving their ideal family were fifty-fifty.

Do you think the Watsons and Johnsons are right about their chances?

Waiting for a girl

Some couples with children feel that their family is not complete until they have at least one boy and one girl. Some long for a boy. Others long for a girl.

Suppose that a couple who want a daughter decide to continue having children until a girl is born. They could be 'lucky' with their first child turning out to be a girl, or they may have a long line of boys before eventually having a daughter. A number of questions arise, the answers to which might well be of interest to the parents. For example, if boys and girls are equally likely, how many children should they expect to have before their family is complete? That is, what is the average size of families who continue having children until a girl is born? What is the most likely size for their family? What is the probability that they will have more than four children?

Waiting for a six

In some board games, players can join in the game only when they obtain a six on the roll of a die. Several questions spring to mind here. First, on average, how many times will a player have to roll the die in order to start? Secondly, what is the most likely number of rolls needed? And what is the chance that a player will still be waiting to join in after 10 rolls, or after 20 rolls?

Write down your ideas about these questions. What does your intuition suggest to you?

Collecting a complete set of musicians

Some time ago, a certain cereal manufacturer offered eight different toy musicians as gifts in packets of a particular popular breakfast cereal. Each packet contained one musician only, but there was no way of knowing which it contained without opening the packet. How many packets might you expect to have to buy in order to acquire a complete set of musicians? That is, what is the average number of packets that a family might have to buy to acquire a complete set? Write down your 'intuitive' answer to this question.

Coinciding birthdays

Suppose that each of the 24 children (no twins) in a (small) class decides to give a party on their birthday. What do you think the chances are that at least two of the children will need to hold a joint party as their birthdays are on the same day of the year?

Summary of Section 1

In this section, the idea of a probability has been introduced and you have been invited to propose answers to a selection of problems involving chance events. In the next section, you will be invited to explore some of these problems using your computer and to make some further hypotheses using your results.

2 Exploring the problems

To study this section, you will need access to your computer, together with the statistics software and Computer Book D.

In Section 1, the probability of an event was defined to be the proportion of the number of times that an event occurs 'in the long run'. So, in theory, we could estimate the probability of an event by carrying out a long sequence of identical experiments and observing the results. In Activity 1.1, you tossed a coin 30 times and noted the proportion of times that it landed heads up. It was observed that the proportion of heads fluctuated as the number of tosses increased, but that it seemed to be settling down to some constant value. It looked as though this value might be $\frac{1}{2}$, but with only 30 tosses, we could not be absolutely sure of this: we really need to carry out a much longer sequence of tosses. No doubt you found that tossing a coin and recording the outcome soon became tedious, even for as few as 30 tosses, so you certainly would not want to toss a coin 300 times! Fortunately, the computer can help you with this sort of task. Of course, the computer does not actually toss a coin, but instead it generates outcomes according to a random procedure. In the case of tossing a coin, it generates sequences of 'heads' and 'tails' which are indistinguishable from the sorts of results you might get if you actually tossed a coin. This sort of alternative to carrying out a real experiment is known as *simulation*.

In this section, you will be using computer simulations, first to investigate further the 'settling down' phenomenon noticed with coin-tossing, and then to explore some of the problems described in Subsection 1.3.

Refer to Computer Book D for the work in this section.

Summary of Section 2

The main purpose of this section has been to encourage you to develop your understanding of the nature of randomness and to compare some of your intuitions from Subsection 1.3 with what happens in practice.
You have had the opportunity to question your intuitions and to start to develop hypotheses about some results. And you have gained experience in using a range of simulations to model situations involving chance.

3 Equally-likely outcomes

In Subsection 1.2, the probability of an event was defined as the long-run proportion of the number of times that the event occurs. It was noted that in some situations involving chance, the probability of a particular outcome can be calculated without recourse to carrying out a sequence of trials or to collecting masses of data. This is the case whenever it is clear that the different possible outcomes are all equally likely to occur. For example, when tossing a coin, the coin seems as likely to land heads up as tails up; and when a die is rolled, each of the six faces seems equally likely to come up.

Several of the problems from Subsection 1.3 can be tackled using the idea of equally-likely outcomes – for example, *The three-card game*, *D'Alembert's heads*, *Balanced families*, *Galileo and the three-dice problem* and *The Chevalier de Méré: sixes and double-sixes.* The aim of this section is to introduce some basic rules for calculating probabilities and to use them to tackle each of these problems in turn.

Activity 3.1 The language of probability: outcomes and events

Before beginning work on these problems, consider briefly the way in which the two words *outcome* and *event* have been used in this chapter. These words have not been used interchangeably: care has been taken over when each is used. We have spoken of 'the possible outcomes of an experiment' and of various 'events associated with an experiment'. The best way for you to sort out the distinction between 'outcomes' and 'events' is to consider some examples.

(a) Suppose that an experiment involves rolling a die with faces numbered from 1 to 6 and noting the score on the uppermost face. First write down all the possible outcomes of the experiment: there are six of them. Then write down at least three events associated with the experiment (there are many possibilities for these). If you are not sure of the difference between 'outcomes' and 'events', then look back at some of the examples and activities in Subsection 1.2 to help you make your lists.

(b) Another experiment involves drawing a card from a well-shuffled pack of 52 playing cards and noting which card it is. There are 52 possible outcomes of the experiment. What are they? Write down three events associated with the experiment.

The distinction between outcomes and events will be useful in this section when developing some basic rules for calculating probabilities. Try to express in your own words what you understand to be the difference between 'outcomes' and 'events'.

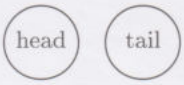
Comment

Consider first experiment (a), in which a die is rolled and the number on the uppermost face is noted. This number is the *outcome* of the experiment; it describes *precisely* what occurs when the specific experiment is run. In this case, there are six possible outcomes of the experiment: the numbers 1 to 6. An *event* can be *any* happening associated with the experiment though it does, of course, include all the outcomes as possible events. Some examples of events associated with this experiment are 'obtaining an even number', 'obtaining a number greater than 4', 'obtaining a 6' and 'obtaining a multiple of 3'.

Similarly, in experiment (b), when a card is drawn from a pack of 52 playing cards and its suit and number are noted, there are 52 possible outcomes: ace of hearts, two of hearts, three of hearts, ..., king of spades. Examples of events associated with this experiment, include 'obtaining a heart', 'obtaining an ace', 'obtaining a black card' and 'obtaining a red queen'.

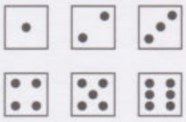
In general, an outcome is the precise result of an experiment (such as getting a six when a die is rolled), whereas an event is *any* happening associated with the experiment: it may be one of the possible outcomes of the experiment, such as 'obtaining a six', or it may be a more general happening, such as 'getting an even number'. One part of probability theory involves developing ways of calculating the probability of any event given only the probability of the possible outcomes.

3.1 Counting problems



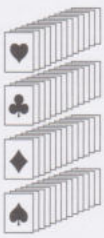
The idea of counting equally-likely outcomes was used in Subsection 1.2 to find the probability of each of a number of events. For example, there are two outcomes when a coin is tossed: heads and tails. Assuming that these are equally likely to occur, each outcome will occur half the time in the long run, so

$$P(\text{head}) = P(\text{tail}) = \tfrac{1}{2}.$$



There are six outcomes when a die is rolled – the faces are numbered 1 to 6. It has been known for some gamblers to cheat by 'loading' their dice so that, when rolled, some faces are more likely to come up than others. However, assuming that a die is not loaded, the six possible outcomes, 1 to 6, are equally likely, so

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \tfrac{1}{6}.$$



Similarly, if a card is drawn from a pack of 52 playing cards, there are 52 possible outcomes and these are all equally likely, so, for example,

$$P(\text{ace of spades}) = \tfrac{1}{52}, \quad P(\text{seven of hearts}) = \tfrac{1}{52}.$$

In general, if an experiment (tossing a coin, rolling a die, picking a card, etc.) has N possible outcomes and these are all equally likely, then for any particular outcome, the probability that it occurs is $1/N$; that is,

$$P(\text{particular outcome}) = \frac{1}{N}.$$

To find the probability that an even score is obtained when a die is rolled, we also counted the number of outcomes that give an even number, and

hence found the proportion of outcomes which give an even score. Three of the six possible scores are even (2, 4 and 6), so

$$P(\text{even score}) = \tfrac{3}{6} = \tfrac{1}{2}.$$

Similarly, to find the probability of drawing an ace from a pack of 52 cards, we counted the number of aces (4) and hence found the proportion of outcomes that give an ace:

$$P(\text{ace}) = \tfrac{4}{52} = \tfrac{1}{13}.$$

These are examples of the following general result.

> If an experiment has N equally-likely possible outcomes, and $n(E)$ is the number of these outcomes that result in an event E occurring, then
>
> $$P(E) = \frac{n(E)}{N}; \tag{3.1}$$
>
> that is, $P(E)$ is equal to the number of outcomes for which the event E occurs divided by the total number of possible outcomes.

This result can be used to answer several of the questions which were posed in Section 1. *The three-card game*, *D'Alembert's heads*, *Balanced families* and *Galileo and the three-dice problem* can all be tackled by counting equally-likely outcomes and using formula (3.1). We shall begin by looking at the problem of the three-card game, which is the simplest of these to solve.

The three-card game

In this game, three cards are put into a hat: one of the cards is white on both sides, one is red on both sides, and the third is white on one side and red on the other. One of the three cards is drawn at random from the hat and placed flat on a table. If the upper side of the card on the table is red, the fairground entertainer offers to bet you 50p that the other side is red, since, as he says, 'This isn't the white-white card, so it must be one of the other two'. Would you accept the wager?

At first sight, the bet might seem a fair one: there are two possibilities: either the card is the red-red one or it is the red-white one. However, the fairground entertainer is no fool – his trick is a good steady earner for him. To see why this is so, we need to calculate the probability that the other side of the card is red. And to do this, we must first identify the equally-likely outcomes involved in the situation. The side showing could be any one of the *three* red sides and, since the card is selected at random and placed on the table, each of these three sides is equally likely to be the one showing. In one case the other side is white, and in the other two cases the other side is red, so

$$P(\text{other side is red}) = \tfrac{2}{3}.$$

The crucial point here is that the equally-likely outcomes are the sides not the cards.

If you find this difficult to understand, then imagine that the sides of the cards are numbered: R1 and R2 on the red-red card, W1 and W2 on the white-white card, and R3 and W3 on the red-white card. Suppose that a card is selected at random and placed on the table, and the side showing is

red. Since the card is selected at random, no side is more likely than any other side to come up – R1, R2 and R3 are equally likely. If the side showing is R1, then the other side is R2; if it is R2, then the other side is R1; and if it is R3, then the other side is W3 (see Figure 3.1).

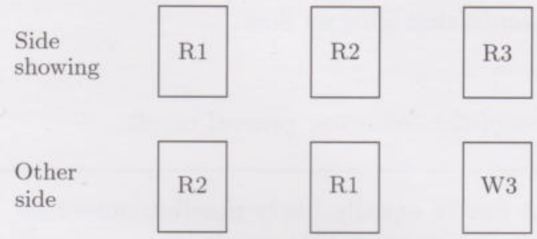


Figure 3.1 The other sides

In two of these three cases, the other side is red and hence the probability that the other side is red is $\frac{2}{3}$. So, in the long run, the entertainer will win approximately $\frac{2}{3}$ of his wagers, and hence make a good profit.

Many people find this result difficult to believe even after following the above argument. If you are not convinced, then try an experiment. Take three pieces of card and label the sides 'red' or 'white' so that one card is labelled red on both sides, one is labelled white on both sides and the third is labelled red on one side and white on the other. Then carry out a sequence of trials. Place the three pieces of card in a bag or a hat, remove a card at random and place it flat on a table without looking at it first. If the side showing is labelled red, then note down the label on the other side – red or white. (If the side showing is labelled white, return the card to the bag and start again.) Repeat this to obtain a sequence of results. Estimate for yourself the proportion of the time that the other side is labelled red.

Activity 3.2 The three-card game

If the side of the card showing on the table is white, the fairground entertainer offers to bet you 50p that the other side is white. Is this a fair bet?

A solution is given on page 55.

D'Alembert's heads

D'Alembert argued that when a coin is tossed twice, the probability that at least one head is obtained is $\frac{2}{3}$. One set of results obtained from running the simulation in Activity 2.4 (in Computer Book D) is shown in Figure 3.2.

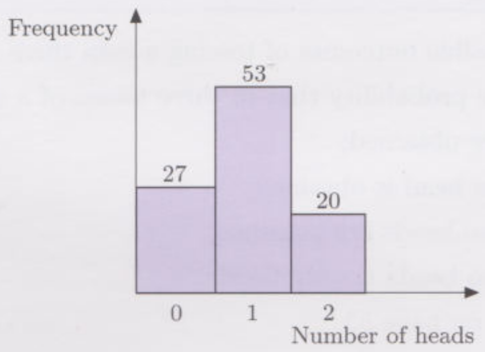


Figure 3.2 The results of a simulation

In the simulation, each trial consisted of tossing a coin twice and noting the number of heads obtained (0, 1 or 2); 100 trials were carried out, and a frequency diagram displayed for the number of heads obtained in the trials. Figure 3.2 shows that in this simulation, at least one head was obtained in 73 ($= 53 + 20$) out of the 100 trials. The proportion of trials in which at least one head was obtained is 0.73 – somewhat greater than $\frac{2}{3}$. Did you obtain similar results? In each of your simulations, was the proportion of trials in which at least one head was obtained greater than $\frac{2}{3}$?

In four further simulations that we carried out, the proportions obtained were 0.68, 0.78, 0.71, 0.75. If these results are typical, then they suggest that the probability of obtaining at least one head in two tosses of a coin is greater than $\frac{2}{3}$ and d'Alembert was wrong!

As with the three-card game, the key step in investigating the problem of *D'Alembert's heads* is to identify the equally-likely outcomes involved in the situation. You are asked to do this in the next activity.

Activity 3.3 Two tosses of a coin

List all the possible outcomes of tossing a coin twice, using h to represent a head and t for a tail. How many outcomes have you listed, and are they all equally likely? What do you make the probability of obtaining at least one head in two tosses of a coin?

Comment

It is important in this activity to distinguish between the results of the first and second tosses. Writing the results of the tosses in the order in which they occur, we obtain *four* possible outcomes:

$$hh, \quad ht, \quad th, \quad tt,$$

where, for example, ht means the first toss results in a head and the second in a tail.

Since a head and a tail are equally likely to occur on each toss, these four outcomes are equally likely. In three of these four outcomes – hh, ht and th – at least one head occurs so, using formula (3.1), the probability of obtaining at least one head in two tosses of a coin is $\frac{3}{4}$. As we suspected, d'Alembert was wrong!

Look back at d'Alembert's argument. Can you see where he went wrong? He identified three events, but these were not equally likely; he did not identify the *equally-likely* outcomes of tossing a coin twice.

Activity 3.4 Three tosses of a coin

(a) List all the possible outcomes of tossing a coin three times.

(b) Write down the probability that in three tosses of a coin:

(i) no heads are obtained;

(ii) at least one head is obtained;

(iii) at least two heads are obtained;

(iv) exactly two heads are obtained.

A solution is given on page 55.

Comment

Note that it is important to be systematic when listing the possible outcomes of an experiment, so as to avoid missing any out.

Did you notice that the two probabilities calculated in parts (b)(i) and (b)(ii) of this activity summed to 1? Since either there are no heads in three tosses or there is at least one head, one or other of these two events is certain to occur. Hence the probability that one or other of the events occurs is equal to 1. Since the two events cannot occur simultaneously, their separate probabilities must add up to 1. This is an example of the following useful rule for probabilities.

If E is an event and not-E is the opposite event (that E does not occur), then

$$P(E) + P(\text{not-}E) = 1,$$

or, equivalently,

$$P(E) = 1 - P(\text{not-}E). \tag{3.2}$$

This rule is a particularly useful one in problems where it is easier to calculate the probability that a particular event does *not* occur than it is to calculate directly the probability that it does. For example, the rule could be used to calculate the probability of at least one head in three tosses of a coin without counting all the possible ways in which at least one head can occur. Using (3.2),

$$\begin{aligned} P(\text{at least one head}) &= 1 - P(\text{no heads}) \\ &= 1 - \tfrac{1}{8} \\ &= \tfrac{7}{8}, \end{aligned}$$

which is the answer you obtained in Activity 3.4 by counting outcomes.

Balanced families

In the 18th century, it was observed that patterns in the sequences formed by the sexes (male and female) of successive births in city hospitals were not unlike the patterns of heads and tails resulting from successive tosses

of a coin. This suggests a possible simple model for births: the probability that the next child born is a girl is $\frac{1}{2}$, and the probability that it is a boy is $\frac{1}{2}$, and these probabilities are the same whatever the sex of children born previously.

If we accept this model, then questions about family patterns can be tackled by methods which have been developed to answer problems about coin-tossing. For example, if we make the analogy between a family of two girls and getting two heads in two tosses of a coin, then the proportion of families of size two consisting of two girls should be approximately equal to the probability of getting two heads in two tosses of a coin. This model can be used to tackle questions like *Balanced families*, for instance.

In fact, as long ago as the middle of the 17th century, John Graunt discussed the sex ratio: the records of christenings that he examined seemed to suggest that rather more boys than girls were born. And in the 18th century, Abraham de Moivre remarked on observations made by Nicholas Bernoulli (1687–1759) as follows.

> Mr Bernoulli collects from Tables of Observations continued for 82 years, that is from A.D. 1629 to 1711, that the number of Births in London was, at a medium, about 14 000 yearly: and likewise, that the number of Males to that of Females ... is nearly as 18 to 17. But he thinks it the greatest weakness to draw any Argument from this against the Influence of Chance in the production of the two sexes. For, says he, 'Let 14 000 Dice, each having 35 faces, 18 white and 17 black, be thrown up, and it is great Odds that the numbers of white and black faces shall come as near, or nearer, to each other, as the numbers of Boys and Girls do in the Tables.'

This quotation is taken from *The Doctrine of Chances* by Abraham de Moivre, which was published in London in 1756.

More recent investigations have confirmed that the proportion of babies born that are boys is slightly greater than $\frac{1}{2}$. Nevertheless, the simple model that has been suggested can be used to obtain approximate results for problems such as *Balanced families*; and, as you will see later in this section, it is not difficult to modify the model to take account of the slight imbalance between male and female births. But in the next activity, you should assume the simple model; that is, you should assume that each child born is equally likely to be a girl or a boy.

Activity 3.5 Balanced families

(a) (i) List all the possible patterns of families of two children, using G to represent a girl and B for a boy. Take care to distinguish between the first born and the second born.

(ii) The Watsons' ideal family is one girl and one boy. What is the probability that they will achieve their ideal family?

(b) (i) List all the possible patterns of families of four children. How many different patterns are there?

(ii) The Johnsons' ideal family is two girls and two boys. What is the probability that they will achieve their ideal family?

(c) Are the Watsons and the Johnsons correct to believe that their chances of achieving their ideal families are both fifty-fifty?

A solution is given on page 55.

Galileo and the three-dice problem

When Galileo wrote down his ideas about the problem of why, when three dice are rolled, a total of 10 seemed to be 'more advantageous' than a total of 9, he started by counting the total number of different possible outcomes. He began as follows.

Source: *Considerazione sopra il Giuoco dei Dadi* (Thoughts about Dice Games).

> ... since a die has six faces, and when thrown it can equally well fall on any one of these, only six throws can be made with it, each different from all the others. But if together with the first die we throw a second, which also has six faces, we can make 36 throws each different from all the others, since each face of the first die can be combined with each face of the second ....

For tossing two coins, it is important to distinguish between the first coin and the second. In the same way, it is important to distinguish between the score on the first die and the score on the second; so, for example, 5 on the first die and 4 on the second is a different outcome from 4 on the first die and 5 on the second. To remind yourself that these are different outcomes, it is sometimes helpful to imagine the dice being different colours; then the fact that they are different outcomes becomes clearer. The argument that Galileo used to calculate the number of possible outcomes when two dice are rolled can be extended to three dice.

Activity 3.6 Rolling three dice

How many different possible outcomes are there when three dice are rolled? (Remember to distinguish between the three dice when counting outcomes – for instance, by imagining that they are different colours.)

Comment

For each of the 36 outcomes for the first two dice, there are 6 possible outcomes for the third die. Hence there are $36 \times 6 = 216$ different possible outcomes when three dice are rolled, and these are all equally likely.

Activity 3.7 Totals of 9 and 10

Having calculated that the total number of possible outcomes is 216, Galileo produced a table showing the number of possible outcomes which give totals of 3, 4, 5, 6, 7, 8, 9 and 10. (He also observed that the totals for 11 to 18 were symmetrical with these; for example, the number of ways of getting a total of 11 is equal to the number of ways of getting a total of 10, and the number of ways of getting a total of 12 is the same as the number of ways of getting a total of 9. We shall not be checking all his results!)

(a) Galileo's table indicated that the number of possible outcomes giving a total of 9 is 25. However, as already noted, there are only 6 different sets of scores of three dice which add up to 9: (621), (531), (522), (441), (432), (333). How do you account for this difference?

(b) List all the possible outcomes that lead to a total score of 9, and hence confirm that Galileo's figure of 25 is correct. What is the probability of obtaining a total score of 9 when three dice are rolled?

(c) Count the possible outcomes which give a total of 10, and hence find the probability of obtaining a total score of 10 when three dice are rolled.

A solution is given on page 56.

In this activity, you found that the probability of a total score of 9 is $\frac{25}{216} \simeq 0.116$ and the probability of a total score of 10 is $\frac{27}{216} = 0.125$. The difference between these probabilities is only $\frac{2}{216}$ or $\frac{1}{108}$. Perhaps the most interesting thing about this result is that the person (the Grand Duke?) who asked Galileo to look at the problem had gambled often enough to detect the effect of so small a difference in probabilities. This suggests how much gambling there must have been among some sections of Italian society at that time.

3.2 Independence and the multiplication rule

The idea of independent events has been used implicitly in many of the examples discussed in Subsection 3.1. For example, it was assumed that the score obtained on rolling a die has no influence on the score obtained on rolling the same die a second time or on rolling a second die. Whether a coin lands heads up when tossed is unaffected by whether it landed heads up the last time it was tossed. And when modelling the births of boys and girls, it was assumed that whether a baby born is a girl or a boy is unaffected by whether babies born earlier were girls or boys. The independence of two events can be defined as follows.

> Two events are **independent** of each other if the occurrence (or not) of one is not influenced by whether or not the other occurs.

The calculation of probabilities involving independent events is often straightforward.

Suppose, for instance, that we want to find the probability that when two coins are tossed, they both land heads up. (You can think of the two coins being tossed together or one after the other – it does not matter which, the result below holds in either situation.) Whether or not one coin lands heads up is clearly not influenced by whether the other lands heads up: the event 'the first coin lands heads up' is independent of the event 'the second coin lands heads up'. Each coin has probability $\frac{1}{2}$ of landing heads up. So, in the long run, the first coin will land heads up half of the time and the second coin will land heads up on half of the occasions that the first coin lands heads up. Therefore, in the long run, the overall proportion of the time that both coins land heads up is

$$\tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{4},$$

that is,

$$P(\text{both coins heads up}) = P(\text{first coin heads up}) \times P(\text{second coin heads up}).$$

Similarly, if a coin and a die are thrown together, then in the long run the coin will land heads up half the time and the die will show a six on $\frac{1}{6}$ of the occasions that the coin lands heads up. So, in the long run, the proportion of the time that a head and a six are obtained together is

$$\tfrac{1}{2} \times \tfrac{1}{6} = \tfrac{1}{12},$$

that is,

$$P(\text{head and six}) = P(\text{head}) \times P(\text{six}).$$

These two examples illustrate the multiplication rule for independent events, which can be stated formally as follows.

Multiplication rule for independent events
If E and F are independent events, then

$$P(E \text{ and } F) = P(E) \times P(F). \qquad (3.3)$$

This rule can be used to calculate some of the probabilities that were found in Subsection 3.1 simply by counting. For example, when a coin is tossed twice, the probability of a tail on the first toss is $\frac{1}{2}$ and the probability of a tail on the second toss is $\frac{1}{2}$, so the probability of tails on both tosses, that is, no heads, is

$$P(\text{no heads}) = \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{4}.$$

Hence, using (3.2),

$$\begin{aligned} P(\text{at least one head}) &= 1 - P(\text{no heads}) \\ &= 1 - \tfrac{1}{4} \\ &= \tfrac{3}{4}, \end{aligned}$$

as obtained by counting outcomes in Activity 3.3.

The multiplication rule extends to three or more independent events in an obvious way: the probability that all the events occur is obtained by multiplying together the probabilities of the separate events. You will need to use this in the next activity.

Activity 3.8 Families

(a) Use the multiplication rule to find the probability that the first three children in a family are all girls.

(b) What is the probability that a family of three will contain at least one boy?

(c) In the first half of the 20th century, Mr and Mrs Grover C. Jones of Peterson, West Virginia, had an all-son family of fifteen sons.

(i) What is the probability that a family of fifteen children will all be boys?

(ii) What is the probability that a family of fifteen children will all be the same sex?

A solution is given on page 56.

Activity 3.9 Bernoulli's families

Nicholas Bernoulli suggested that the births of boys and girls could be modelled by the rolling of a die with 35 faces, 18 of which represent a boy and 17 of which represent a girl.

(a) According to this model, what is the probability that a baby born is a girl?

(b) What is the probability that a family of three children will all be girls?

(c) What is the probability that a family of three children will contain at least one boy?

A solution is given on page 56.

In a similar way, since the results of successive rolls of a die are independent and the scores on different dice are independent, the multiplication rule can be used to answer questions about successive rolls of a die, or about rolls of two, three or four dice. For example, the probability of obtaining two sixes in two rolls of a die is found by multiplying the probability that the first roll gives a six by the probability that the second roll gives a six:

$$P(\text{two sixes in two rolls}) = \tfrac{1}{6} \times \tfrac{1}{6} = \tfrac{1}{36}.$$

Use the multiplication rule to find the probabilities in the next two activities.

Activity 3.10 Hazard

One of the gambling games with which the 16th-century Italian mathematician Girolamo Cardano was very familiar was called *Hazard*. In Italy, this game was played with three dice. In his autobiography, *De Vita Propria Liber* (The Book of My Life), Cardano wrote the following about the total score obtained when three dice are rolled.

In England and some other European countries, Hazard was a game played with two dice.

> To throw in a fair game at Hazards only three spots ... is a natural occurrence and deserves to be so deemed; and even when they come up the same way for a second time, if the throw be repeated. If the third and fourth plays are the same, surely there is occasion for suspicion on the part of a prudent man.

(a) What is the probability that when three dice are rolled, the total score on the three dice is 3 (that is, in Cardano's terminology, the total number of spots uppermost is three)?

(b) Find the probability that a total score of 3 is obtained: (i) in each of two successive rolls of three dice; (ii) in each of three successive rolls of three dice; (iii) in each of four successive rolls of three dice. Do you agree with Cardano that you should be suspicious if a score of 3 occurs three or four times in a row?

A solution is given on page 56.

Activity 3.11 Dice problems

(a) Find the probability of obtaining no sixes in two rolls of a die.

(b) Find the probability of obtaining no sixes in three rolls of a die.

(c) Find the probability of obtaining a double-six when two dice are rolled once.

(d) Find the probability of obtaining two double-sixes in two rolls of a pair of dice.

A solution is given on page 56.

The Chevalier de Méré: sixes and double-sixes

The two simple rules (3.2) and (3.3) are sufficient to tackle the problem which was posed to Blaise Pascal by the Chevalier de Méré. The Chevalier knew that to bet on obtaining at least one six in 4 rolls of a single die was advantageous to him in the long run. He wanted to know why it was not also advantageous to bet on obtaining at least one double-six in 24 rolls of a pair of dice. (Presumably he discovered this the hard way – by experience!)

A useful point to note here is that 'at least' problems are often best tackled 'backwards', that is, using rule (3.2):

$$P(E) = 1 - P(\text{not-}E).$$

That is certainly the case for de Méré's problem. First consider the probability of obtaining at least one six in 4 rolls of a single die:

$$P(\text{at least one six in 4 rolls}) = 1 - P(\text{no sixes in 4 rolls}).$$

The probability of failing to get a six in a single roll is $\frac{5}{6}$ so, using the multiplication rule (3.3),

$$P(\text{no sixes in 4 rolls}) = \left(\tfrac{5}{6}\right)^4.$$

Hence

$$P(\text{at least one six in 4 rolls}) = 1 - \left(\tfrac{5}{6}\right)^4 \simeq 0.518.$$

This is greater than $\frac{1}{2}$, thus confirming that betting on obtaining at least one six in 4 rolls of a single die is advantageous in the long run.

In the next activity, you are asked to work out the probability of obtaining at least one double-six in 24 rolls of a pair of dice: this is the probability that de Méré needed to know.

Activity 3.12 De Méré's problem

(a) What is the probability of failing to obtain a double-six in a single roll of a pair of dice?

(b) Find the probability of failing to obtain any double-sixes in 24 rolls of a pair of dice.

(c) Hence find the probability of obtaining at least one double-six in 24 rolls of a pair of dice. Was the Chevalier correct to suspect that making the second wager was not a good idea?

A solution is given on page 57.

Summary of Section 3

In this section, the concept of equally-likely outcomes has been used to introduce some basic rules for calculating probabilities. These rules have been used to tackle five of the problems described in Subsection 1.3: *The three-card game*, *D'Alembert's heads*, *Balanced families*, *Galileo and the three-dice problem* and *The Chevalier de Méré: sixes and double-sixes.*

Exercises for Section 3

Exercise 3.1 Lucky tickets?

I recently attended a cricket club presentation evening where I was invited to draw a ticket out of a hat in exchange for 50p. To win a prize, the number on the ticket had to end in 0 or 5. Suppose that there were 500 tickets in the hat, numbered from 1 to 500, and that I was the first person to draw a ticket.

(a) How many tickets in the hat had numbers ending in 0 or 5?

(b) What was the probability that I would win a prize?

Exercise 3.2 Tetrahedral dice

Two tetrahedral dice each have faces labelled 1, 2, 3 and 4. The dice are rolled, and a note is made of the number on the face on which each die lands.

A *tetrahedron* is a regular four-sided solid, with each face an identical equilateral triangle. Note that for a tetrahedral die, the score is that on the face on which it lands, as opposed to that on the uppermost face on a cubical die.

(a) Find the probability that the numbers obtained on the dice add up to 5.

(b) Find the probability that the first die lands on a 2 and the second die lands on an odd number.

Exercise 3.3 More about tetrahedral dice

(a) A pair of tetrahedral dice are rolled. What is the probability of obtaining a double-four?

(b) Hence find the probability of failing to obtain a double-four in a single roll of a pair of tetrahedral dice.

(c) Find the probability of failing to obtain any double-fours in six rolls of a pair of tetrahedral dice.

(d) Hence find the probability of obtaining at least one double-four in six rolls of a pair of tetrahedral dice.

4 Waiting for a success

In this section, the ideas introduced in Section 3 will be used to investigate two of the problems described in Subsection 1.3 and explored in the computer section: *Waiting for a six* and *Waiting for a girl.* The first of these two problems is described again below.

Waiting for a six

Suppose that you are playing a board game in which players can join in the game only when they roll a six with a die. Here are some questions about the time (measured in terms of the number of rolls) you might have to wait to join in: those posed in Subsection 1.3 are included.

Question 1: What is the probability that you will be able to join in the game straightaway? That is, what is the probability that you will roll a six at the first attempt?

Question 2: What is the probability that you will join in the game after your second roll of the die? Or after your third roll? Or after five or ten rolls? Are you more likely or less likely to join in the game after your fifth roll than after your tenth roll?

Question 3: When are you most likely to join in the game? That is, what is the most likely number of times you will need to roll the die to get a six?

Question 4: How likely is it that you will still be waiting to join in the game after five rolls of the die, or after ten rolls, or even after twenty rolls?

Question 5: On average, how many times will you have to roll the die in order to obtain a six and so join in the game?

You were invited to explore several of these questions in the computer section, so you should have some ideas about their answers. As we tackle the questions in this section, several important ideas from probability theory will be introduced. First, in Subsection 4.1, the notions of a random variable and a probability distribution are discussed. Then, in Subsection 4.2, the idea of the mean of a probability distribution is introduced; this is the mean value predicted by the probability model.

4.1 *Is a long wait likely?*

In Activity 2.6 (in Computer Book D), the number of rolls of a die needed to obtain a six was simulated: this was repeated 300 times to obtain the lengths of 300 waits, and the results were displayed in a frequency diagram. Figure 4.1 shows one set of results obtained by a member of the course team running the simulation.

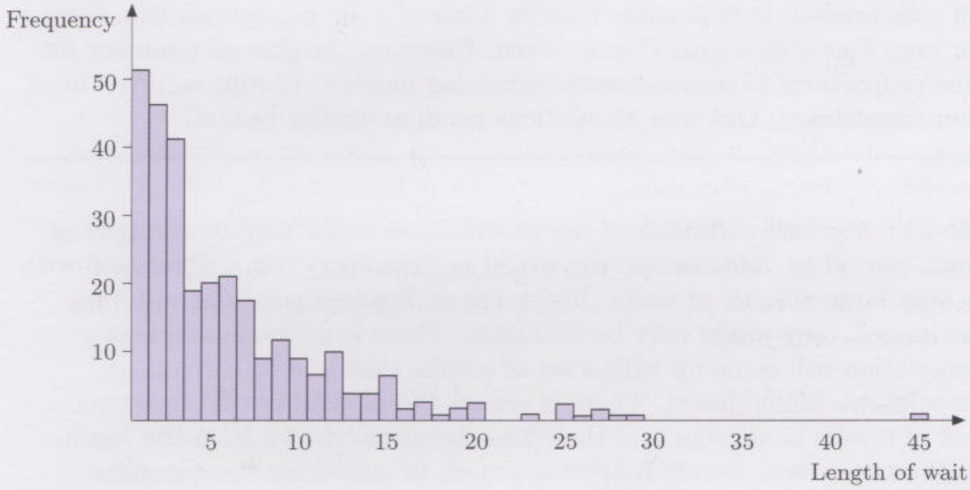


Figure 4.1 The results of a simulation

These results give us some idea of the relative likelihood of different numbers of rolls of a die being needed to obtain a six. But are they typical? Were your results similar? Use your results from Activity 2.6 for simulations with 300 waits, and those in Figure 4.1, to answer the questions in the next activity.

Activity 4.1 How many rolls?

(a) What number of rolls of a die do you think you are most likely to need to obtain a six? That is, at what stage are you most likely to join in the game?

(b) Are you more likely or less likely to need ten rolls of a die than you are to need five rolls?

(c) Summarise in your own words the information contained in Figure 4.1.

Comment

(a) Figure 4.1 shows that just one roll was needed more often than any other number of rolls. That is, it appears that the most likely number of rolls of a die needed to obtain a six is 1. However, notice that one roll was needed only 52 times out of 300, just over one sixth of the time. So it is not all that likely that you will obtain a six with your first roll of the die. Nevertheless, although it may not be very likely, it does seem to be more likely than any of the other possibilities.

(b) From the figure, it looks as though ten rolls is less likely than five rolls: in the simulation, the first six was obtained on the fifth roll 20 times, whereas the first six was obtained on the tenth roll only 9 times.

(c) In general, it appears that you are more likely to need a small number of rolls than you are to need a large number. In fact, a very large number of rolls looks very unlikely. Roughly speaking, the proportion or 'empirical probability' seems to decrease as the number of rolls needed increases.

The simulation was repeated ten times altogether by the course team member: in eight of the simulations, one roll occurred most often, and in the other two simulations, two was the most frequently occurring number of rolls needed. It is possible that for some of your simulations two, three or even four rolls occurred most often. However, the general tendency for the proportions to decrease with increasing numbers of rolls recurred in all our simulations. Did your simulations produce similar results?

To gain accurate estimates of the probabilities of the various numbers of rolls needed to obtain a six, you would need to carry out a simulation with a very large number of waits. Nevertheless, however good the resulting estimates, they would only be estimates. There is no guarantee that a simulation will come up with a set of results that lead to correct conclusions being drawn. There is always the possibility that an atypical set of results may occur and that the conclusions drawn from the results will be incorrect. An alternative approach to answering the questions asked at the beginning of this section is to use the ideas developed in Section 3 to calculate corresponding theoretical results.

It will simplify the arguments and explanations which follow if we represent the number of rolls of a die required to obtain a six by the capital letter X. Note that X is not a fixed number: it may take different values on different occasions – that is a matter of chance. Sometimes X may be 1, on other occasions X may be 2 or 3, or any larger whole number we care to name. In fact, X is an example of a **random variable** – a quantity which may take different values on different occasions.

Having defined the random variable X, from now on we can use the letter X instead of the lengthy phrase 'the number of rolls of a die needed to obtain a six'. For example, the probability of obtaining a six on the first roll of a die can be written as $P(X = 1)$; this is usually read as 'the probability that X is equal to 1'. Similarly, the probability that two rolls are needed to obtain a six can be written as $P(X = 2)$, and so on.

Activity 4.2 Calculating probabilities: theoretical results

Let X be the number of rolls of a single die needed to obtain a six. You will need to use the multiplication rule for independent events – result (3.3) – to work out some of the probabilities asked for below.

(a) Find $P(X = 1)$, the probability that the first roll results in a six.

(b) Find $P(X = 2)$, the probability that the first roll does not result in a six and the second roll does.

(c) Find $P(X = 3)$, the probability that neither of the first two rolls results in a six and the third roll does.

(d) Find $P(X = 4)$.

(e) Suggest a formula for $P(X = j)$, the probability that the first six is obtained on the jth roll, where $j = 1, 2, 3, \ldots$.

Comment

(a) The probability that the first roll results in a six is $\frac{1}{6}$, so

$$P(X = 1) = \tfrac{1}{6}.$$

(b) The probability that the first roll does not result in a six and the second roll does is, using (3.2) and (3.3),

$$P(X = 2) = \tfrac{5}{6} \times \tfrac{1}{6} = \tfrac{5}{36}.$$

the first roll is not a six

the second roll is a six

(c) The probability that neither of the first two rolls results in a six and the third roll does is

$$P(X = 3) = \tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{1}{6} = \tfrac{25}{216}.$$

not a six

not a six

a six

(d) The probability $P(X = 4)$ is found similarly:

$$P(X = 4) = \tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{1}{6} = \tfrac{125}{1296}.$$

the first three rolls are not sixes

the fourth roll is a six

(e) A pattern is forming here: since the first six appears on the jth roll if and only if the first $j - 1$ rolls do not result in a six and the jth roll does, using the multiplication rule, we obtain

$$P(X = j) = \tfrac{5}{6} \times \tfrac{5}{6} \times \cdots \times \tfrac{5}{6} \times \tfrac{1}{6}.$$

the first $j - 1$ rolls are not sixes

the jth roll is a six

That is,

$$P(X = j) = \left(\tfrac{5}{6}\right)^{j-1} \times \tfrac{1}{6}, \quad j = 1, 2, 3, \ldots. \qquad (4.1)$$

The function defined by formula (4.1) is called the **probability function** of X: for each value of j, it gives you the value of the probability $P(X = j)$. For example, if you require $P(X = 3)$, the probability that 3 rolls of a die are needed to obtain a six, then putting $j = 3$ in formula (4.1) gives

$$P(X = 3) = \left(\tfrac{5}{6}\right)^{3-1} \times \tfrac{1}{6} = \left(\tfrac{5}{6}\right)^{2} \times \tfrac{1}{6} = \tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{1}{6}.$$

Formula (4.1) describes completely how likely the different possible values of X are to occur; that is, it describes how the probabilities (which must add up to 1) are distributed among the different possible values, or, using the language of probability, it describes completely the **probability distribution** of X. The probabilities are illustrated in Figure 4.2, which we shall refer to as a probability diagram, for convenience. Compare this probability diagram with the frequency diagram for 1000 simulated waits given in Figure 4.3: the shapes are very similar, but the probability diagram is smoother than the frequency diagram.

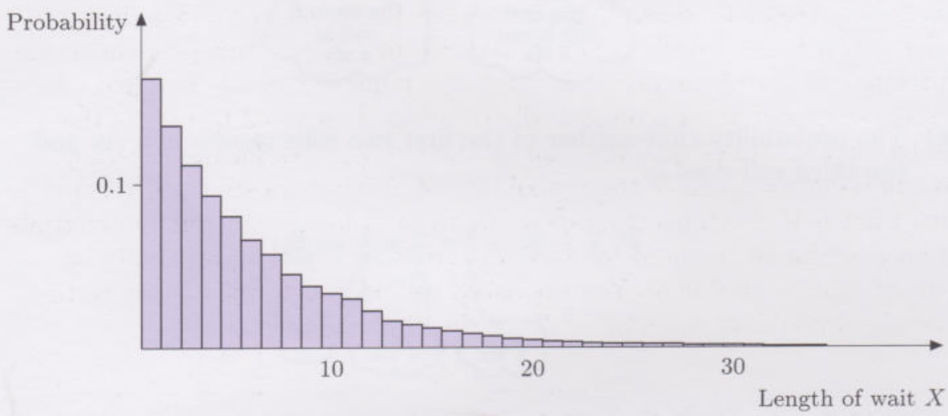


Figure 4.2 The probability distribution of X

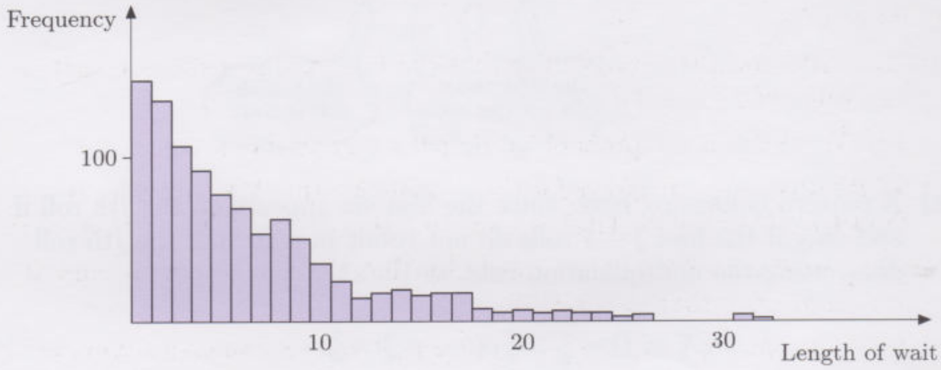


Figure 4.3 The results of simulating 1000 waits

Activity 4.3 Waiting for a six

Use the probability function given by formula (4.1) and illustrated in Figure 4.2 to answer the following questions.

(a) How many rolls of a die are you most likely to require to obtain a six? That is, when are you most likely to join in the board game? What is the probability that you will need this number of rolls?

(b) Find the probability that you will need: (i) exactly five rolls to obtain a six; (ii) exactly ten rolls to obtain a six. Which probability is the greater?

A solution is given on page 57.

Many probability distributions may be used as models for the uncertainty inherent in a wide variety of different situations, and so the most common distributions have been given names. The distribution given by formula (4.1) and illustrated in Figure 4.2 is an example of a **geometric distribution**. It is called a *geometric* distribution because the probabilities $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, ... form a geometric sequence: each probability is obtained from the previous one by multiplying it by a fixed number ($\frac{5}{6}$ in this case).

Suppose that we regard obtaining a six as a success, and rolling a die once as a trial; then we can think of X as the number of trials required to obtain a success. In general, in any sequence of trials of an experiment, each of which may result either in 'success' or 'failure', independent of the outcomes of any of the previous trials, the number of trials required to obtain a success has a geometric distribution. It is usual to denote the probability of success in each trial by the letter p. So, for instance, for rolling a die $p = \frac{1}{6}$, since the probability of obtaining a six (a success) is $\frac{1}{6}$. But what is the formula corresponding to (4.1) for X, the number of trials of an experiment required to obtain a success, when the probability of success in each trial is p? You are asked to find this formula, that is, to find the probability function of X, in the next activity.

Activity 4.4 Waiting for a success

A sequence of trials is carried out: in each trial, the probability of a success is p. The random variable X is the number of trials required to obtain a success.

(a) (i) Write down the probability that the first trial is a success, that is, the value of $P(X = 1)$.

(ii) Write down the probability that the first trial is a failure.

(b) Write down an expression for the probability that the first trial is a failure and the second trial is a success, that is, the value of $P(X = 2)$.

(c) Find an expression for the probability that the first success occurs at the third trial, that is, for $P(X = 3)$.

(d) Suggest a formula for the probability that the first success occurs at the jth trial, that is, suggest a formula for $P(X = j)$.

A solution is given on page 57.

The results obtained in this activity are summarised below.

The geometric distribution

If a sequence of trials of an experiment is carried out and the probability of success in each trial is p $(0 < p < 1)$, then X, the number of trials required to obtain a success, has a geometric distribution. The probability function of X is given by

$$P(X = j) = (1 - p)^{j-1}p, \quad j = 1, 2, 3, \ldots. \qquad (4.2)$$

Figure 4.4 shows the probability distribution of X, the number of trials required to obtain a success, for a typical value of p. The basic shape of the probability diagram is the same whatever the particular value of p.

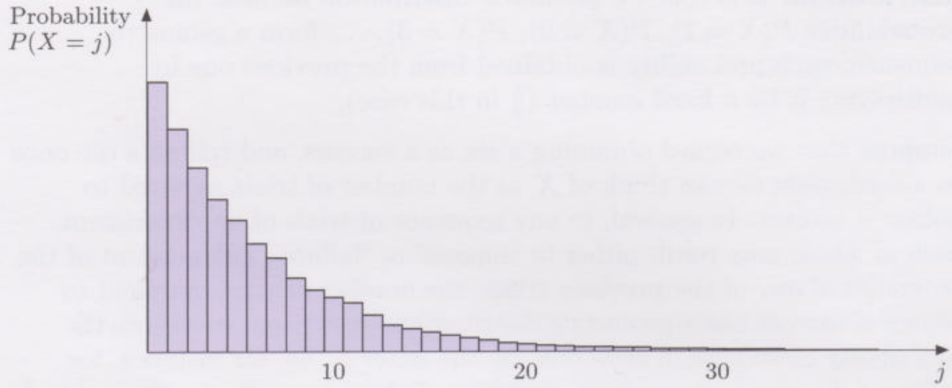


Figure 4.4 A geometric distribution

It is clear from Figure 4.4 and formula (4.2) that, whatever the value of p, the first success is most likely to occur at the first trial, and it is more likely to occur at the second trial than at the third, and so on. So, for example, for a board game where you must roll a six to join in, you are more likely to join in after just one roll of the die than after two, and you are more likely to join in after two rolls than after three, and so on.

Activity 4.5 Waiting for a girl

The Smiths want a daughter, and so decide to continue having children until they have a girl: they will then consider their family to be complete. Suppose that the probability that each child born is a girl is $\frac{1}{2}$, and that the sex of each child does not depend on the sex of previous children. If we regard having a girl as a 'success', then the geometric distribution given in formula (4.2) may be used to model the size of the family they may ultimately have.

(a) What size family are the Smiths most likely to have?

(b) What is the probability that they have only one child?

(c) What is the probability that they have exactly three children?

(d) What is the probability that they have exactly six children?

A solution is given on page 57.

In Activities 4.2 and 4.3, you obtained answers to the first three questions that were posed at the beginning of this section: Questions 1, 2 and 3 on page 34. And in Activity 4.5 you answered some similar questions about the likely family size of a couple who continue their family until they have a daughter.

We shall now turn our attention to Question 4: how likely is it that you will still be waiting to join in a board game after five rolls of the die, or after ten rolls, or even after twenty rolls? No new ideas are needed to tackle this question; you just need to think carefully about the question being asked.

Consider first the probability that you will still be waiting to join in the game after five rolls of the die. This is just the probability that you fail to roll a six on each of the first five rolls so, by the multiplication rule (3.3), it is equal to

$$\tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{5}{6} = \left(\tfrac{5}{6}\right)^5 \simeq 0.402.$$

Notice that you can also think of this as the probability that you will need more than five rolls of the die to obtain a six, since you will need more than five rolls if none of the first five rolls results in a six, and vice versa; the two events are equivalent. So, if X is the number of rolls of the die needed to obtain a six, then we can write

$$P(X > 5) = \left(\tfrac{5}{6}\right)^5;$$

that is, the probability that more than five rolls are needed to obtain a six is $\left(\tfrac{5}{6}\right)^5$.

Activity 4.6 How likely is a long wait?

(a) Find the probability that you will still not have scored a six after ten rolls of the die, that is, find $P(X > 10)$, the probability that you will need more than ten rolls of the die to obtain a six.

(b) Find the probability that you will need at most ten rolls of the die to obtain a six.

(c) Find the probability that you will need more than twenty rolls of the die to obtain a six, that is, work out $P(X > 20)$.

A solution is given on page 57.

Activity 4.7 How likely is a large family?

The Smiths decided to continue having children until they had a daughter (see Activity 4.5).

(a) What is the probability that they will have more than four children?

(b) What is the probability that they will have four or fewer children?

A solution is given on page 58.

4.2 How long is an average wait?

The final question posed at the beginning of this section (page 34) asked 'How many times, on average, will you have to roll the die to obtain a six?', or equivalently 'How long will you have to wait, on average, to join in the board game?'. This is one of the questions that you were invited to explore in the computer section. In Activity 2.6 in Computer Book D, you ran a series of simulations, each involving 300 waits. Part of the output of each simulation was the average length of the waits. Our ten simulations produced the following average waits.

6.2 6.3 6.0 6.3 5.6 6.1 6.3 6.4 6.1 5.8

The average wait varied from one simulation to another because of the nature of the model: it is a model for the uncertainty involved in rolling a die. So how long should you expect to wait, on average, to join in the game? What does the *model* predict for the average wait? The ten values above are all estimates of this mean wait. Some of these values are less than 6 and some are greater than 6, but it looks as though the mean wait predicted by the model should be somewhere close to 6. But how can we find its exact value?

In Subsection 1.2, the probability that an event occurs was defined to be the proportion of the time 'in the long run' that the event occurs. This suggests a possible definition for the mean wait predicted by the model: it is the long-run average wait. If we simulate a long sequence of waits and calculate the average wait as each wait is simulated, then these averages should settle down to the mean wait that we are seeking.

Suppose that in a simulation of a long sequence of waits, a wait of length 1 occurs f_1 times, a wait of length 2 occurs f_2 times, and so on. Then the mean wait is given by

$$\text{average wait} = \frac{1}{n}(1 \times f_1 + 2 \times f_2 + 3 \times f_3 + \cdots),$$

where n is the total number of waits in the sequence (that is, $n = f_1 + f_2 + f_3 + \cdots$).

The mean $\overline{x}$ of a sample of n observations is given by the formula

$$\overline{x} = \frac{1}{n}(x_1 f_1 + x_2 f_2 + \cdots + x_k f_k),$$

where $x_1, x_2, \ldots, x_k$ are the different values observed in the sample, and $f_1, f_2, \ldots, f_k$ are their frequencies.

We can rewrite this as

$$\text{average wait} = 1 \times \frac{f_1}{n} + 2 \times \frac{f_2}{n} + 3 \times \frac{f_3}{n} + \cdots.$$

The mean wait predicted by the model is the long-run value of this average wait. However, f_1/n is the proportion of waits that are of length 1, and for a long sequence of waits this proportion will be approximately equal to the probability that one roll is required to obtain a six; that is, $P(X = 1)$. Similarly, f_2/n will be approximately equal to $P(X = 2)$, and so on. Hence, as we take longer and longer sequences of waits, the average wait will settle down to

$$1 \times P(X = 1) + 2 \times P(X = 2) + 3 \times P(X = 3) + \cdots.$$

So we have the result

$$\text{mean wait} = \sum_{j=1}^{\infty} j \times P(X = j). \qquad (4.3)$$

That is, the mean wait is equal to the sum of the products $j \times P(X = j)$.

The mean predicted by the model is sometimes referred to as the *mean of the probability distribution* or the *mean of the random variable* X, and is denoted by the Greek lower-case letter μ. The mean of a probability distribution is an important idea in probability and statistics, and one to which we shall return in Chapter D2.

The letter μ is pronounced 'mu'.

In general, the mean μ of a random variable X is defined to be

$$\mu = \sum_j j \times P(X = j), \tag{4.4}$$

where the summation is over all values j which X can take (that is, for which $P(X = j) > 0$). You will not be expected to calculate means of probability distributions. In this block, we are interested in the results themselves rather than in the algebra involved in their calculation, so we shall not take you through the details.

You may already have an idea about a formula for the mean of a geometric distribution. In Activity 2.7 in Computer Book D, you were invited to investigate the mean wait by finding the average wait for a number of simulations. For *Waiting for a six*, p, the probability of success at each trial – that is, of obtaining a six with each roll of the die – is $\frac{1}{6}$. We have already observed from the results of some simulations that it looks as though the mean wait is about 6. In Activity 2.7, you were also asked to investigate the mean waits for other values of p: $p = \frac{1}{2}$, $p = \frac{1}{5}$, $p = 0.4$ and some values of your own choosing. No doubt you discovered that for $p = \frac{1}{2}$ the mean wait is about 2, for $p = \frac{1}{5}$ the mean wait is about 5, and for $p = 0.4$ the mean wait is about 2.5. This suggests that, in general, the mean wait is $1/p$.

Using the definition of the mean wait (4.3) and result (4.2), which gives the probability $P(X = j)$ for a geometric distribution, it can be shown that the mean wait is indeed equal to $1/p$. This result is stated in the box below.

The mean of a geometric distribution

If a sequence of trials is carried out and the probability of success in each trial is p, then the mean number of trials required to obtain a success is $1/p$.

This result is derived in an appendix to this chapter.

Use this result to answer the questions in the next activity.

Activity 4.8 Mean waiting times

(a) How many times, on average, will you have to roll a die to obtain a six?

(b) What size family, on average, will couples like the Smiths have in order to get the daughter they long for? (See Activity 4.5.)

(c) Tom hits the bull's-eye on a darts board on roughly $\frac{2}{9}$ of his attempts. How many darts does he need to throw on average to hit the bull's-eye?

A solution is given on page 58.

Summary of Section 4

In this section, the problems *Waiting for a six* and *Waiting for a girl* have been tackled. In the course of investigating these problems, the ideas of a random variable and of a probability distribution and its mean have been introduced. A probability distribution called the geometric distribution has been discussed; this is the probability distribution of the number of trials of an experiment needed to achieve a success. A number of results were derived, and a formula was given for the mean number of trials needed to achieve a success.

Exercises for Section 4

Exercise 4.1 Bernoulli's families

Nicholas Bernoulli suggested that the sex of a child at birth could be modelled by rolling a die with 35 faces, 18 faces representing a boy and the other 17 a girl.

(a) According to this model, what is the probability that a couple such as the Smiths will have four children? (See Activity 4.5, *Waiting for a girl*.)

(b) What is the probability that they will have more than four children?

(c) What is the average family size of couples like the Smiths who continue having children until they have a daughter?

Exercise 4.2 Waiting for the jackpot

In Activity 1.4, you found that if you buy one ticket in the British National Lottery, then the probability of winning a share of the jackpot is $\frac{1}{13\,983\,816}$.

(a) Approximately how many years on average do people who buy one ticket a week have to wait to win a share of the jackpot?

(b) If you buy one ticket a week, what is the approximate probability that you will not yet have won a share of the jackpot after 50 years?

5 Outstanding problems

We have now obtained solutions to all but two of the problems described in Subsection 1.3. In this section, the remaining two problems – *Collecting a complete set of musicians* and *Coinciding birthdays* – will be tackled using some of the ideas and results of the previous sections. You should find that working through these problems will help you to consolidate your understanding of the main ideas and results of this chapter. You may also find the answers interesting and surprising!

5.1 Collecting a complete set of musicians

A cereal manufacturer is giving away a toy musician in each packet of a certain popular breakfast cereal. There are eight different musicians, but there is no way of knowing which musician is inside any particular packet without opening it. The question here is: 'How many packets, on average, will you have to buy to collect a complete set of musicians?'

Activity 5.1 Modelling assumptions

In Activity 2.8 in Computer Book D, you simulated collecting a set of musicians, using the computer. The simulation is based on the assumption that each packet is equally likely to contain any one of the eight different musicians available. What assumptions are we making about the distribution of musicians in the packets of cereal by using the simulation?

Comment

One assumption is that there are equal numbers of the eight musicians available. A second is that no musicians either predominate or are missing from consignments delivered to a particular shop or a particular area.

Figure 5.1 shows the results obtained for one such simulation carried out by a course team member. The musicians are numbered from 1 to 8.

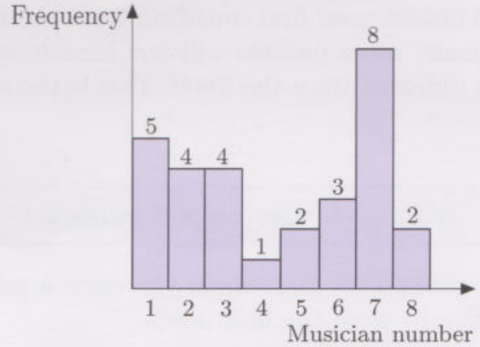


Figure 5.1 The results of one simulation (number of packets = 29)

In this particular simulation, by the time musician number 4 was obtained (the last one needed to complete the set), there were two or more of each of the other seven musicians and no fewer than eight of musician number 7. Altogether, 29 packets were required to complete the set. That seems a lot! Is it typical, or was it simply a run of bad luck? Did you obtain similar results? When the course team member ran the simulation a further nine times, the following numbers of packets were obtained.

42 59 15 20 27 21 23 18 33

The average of the values obtained from these ten simulations is 28.7, so judged on the basis of this evidence, 29 is not an unusually large number of packets. This average, 28.7, is one estimate of the average number of packets required to obtain a complete set of eight musicians.

Activity 5.2 Estimating the mean

Looking at all the results given above, do you think an estimate of 28.7 for the mean number of packets is a good one? Do you think it is likely to be close to the 'true' mean – that is, the mean predicted by the model? Or might it differ quite a lot from this 'true' mean?

Comment

It is possible that 28.7 is a good estimate. However, the results were very variable: the smallest number of packets needed was 15, while in one simulation 59 were required. If only one of the values had been different, then the estimate could have been much larger or much smaller; for instance, if there had been another 15 instead of the 59, then the estimate would have been only 24.3; or if there had been another 59 instead of the 15, then the estimate would have risen to 33.1. Since the results were so variable, far more simulations are needed. So the estimate of 28.7 could be some way from the mean.

The alternative to running simulations in order to estimate the average number of packets required to complete a set is to use the model itself to calculate the mean, that is, to use probability theory. In fact, we can find the mean by making use of some of the ideas from Section 4, 'Waiting for a success'.

Clearly, you will obtain your first musician from the first cereal packet you open. But how many more packets will you have to open to obtain a second musician different from the first? This is the subject of the next activity.

Activity 5.3 Waiting for the second musician

(a) What is the probability that when you open a packet you will find a musician different from the first one?

(b) How many packets, on average, will you need to open to obtain a musician different from the first one?

Comment

(a) Seven out of eight musicians are different from the first one, so the probability that a packet contains a different musician is $\frac{7}{8}$.

(b) As long as you have musicians of only one type, the next packet will contain a different musician with probability $\frac{7}{8}$. So if you regard opening a packet and finding a second musician as a 'success', then the number of packets you will need to open to obtain a 'success' (a second musician) has a geometric distribution. In Subsection 4.2 (page 43), you saw that the mean of a geometric distribution is $1/p$, where p is the probability of success in each trial. So the mean number of *additional* packets needed to obtain a second musician is

$$1/\tfrac{7}{8} = \tfrac{8}{7}.$$

So you need 1 packet to obtain your first musician, and $\frac{8}{7}$ packets, on average, to obtain your second musician. Hence, on average, you will need to open a total of

$$1 + \tfrac{8}{7} \simeq 2.14 \text{ packets}$$

to obtain your first two musicians.

Activity 5.4 Waiting for the third musician

(a) Once you have collected two different musicians, what is the probability that the next packet you open will contain a musician different from the first two?

(b) How many packets, on average, will you need to open to obtain your third different musician?

Comment

(a) Six of the eight musicians are different from the first two, so the probability of finding a different musician in the next packet is $\frac{6}{8}$.

(b) Again, if obtaining a different musician is a 'success', then the average number of packets you will need to open to obtain a 'success' is $1/p$, where $p = P(\text{success})$. So the average number of *additional* packets required to obtain a third different musician is

$$1/\tfrac{6}{8} = \tfrac{8}{6}.$$

And the *total* number of packets required, on average, to obtain your first three different musicians is

$$1 + \tfrac{8}{7} + \tfrac{8}{6} \simeq 3.48.$$

Can you see a pattern developing here?

Activity 5.5 Waiting for the rest of the musicians

(a) Once you have three different musicians, how many additional packets, on average, will you need to open to obtain your fourth different musician?

(b) Once you have four different musicians, how many additional packets, on average, will you need to open to obtain your fifth different musician?

(c) How many additional packets, on average, will you need to open to obtain your sixth, seventh and eighth musicians?

(d) How many packets in total will you need to open, on average, to collect a complete set of eight musicians?

Comment

(a) When you have three different musicians, the probability that the next packet contains a different musician is $\frac{5}{8}$, so the average number of *additional* packets needed to obtain a fourth musician is $1/\frac{5}{8} = \frac{8}{5}$.

(b) Similarly, on average, you will need to open $1/\frac{4}{8} = \frac{8}{4}$ *additional* packets to obtain your fifth different musician.

(c) You will need to open $1/\frac{3}{8} = \frac{8}{3}$ packets, on average, to obtain a sixth different musician, $1/\frac{2}{8} = \frac{8}{2}$ packets to obtain your seventh musician, and $1/\frac{1}{8} = \frac{8}{1}$ packets to obtain your eighth and final musician in the set.

(d) So the mean total number of packets you will need to open to collect a complete set of musicians is

$$1 + \tfrac{8}{7} + \tfrac{8}{6} + \tfrac{8}{5} + \tfrac{8}{4} + \tfrac{8}{3} + \tfrac{8}{2} + \tfrac{8}{1} \simeq 21.74,$$

or approximately 22 packets.

Activity 5.6 Comparing your answers

Compare the theoretical result above with your intuitive answer from Subsection 1.3 and the number you obtained using simulation in Activity 2.8 (in Computer Book D). Are you surprised that the mean number of packets needed to collect a set of size eight is as large as 22? Or were you expecting this sort of result or a larger number after doing the simulations? In what way has simulation helped you to understand what is involved in modelling and solving the problem?

Comment

Most people's intuitive answers are numbers much smaller than 22. Was your value lower than this? If so, then you were probably surprised by the results of your simulations. However, having done the simulations, you were probably not very surprised by the theoretical result – not even if your estimate was as far from the mean predicted using theory as was our estimate of 28.7 (see page 46).

You may also have been surprised by how greatly the number of packets varied from one simulation to the next. This is something you would not have been aware of if you had simply used theory to calculate the mean.

The approach used to find the mean number of packets required to obtain a complete set of eight musicians can be used to find the average size of 'complete' families. Try this in the next activity.

Activity 5.7 When is a family complete?

Some couples do not regard their family as complete until they have at least one boy and one girl, and so decide to continue having children until they have at least one son and at least one daughter. Assuming that a boy and a girl are equally likely, what is the mean size of such families? How does this mean compare with the estimate you obtained in Activity 2.9(b) in Computer Book D, using the simulation software?

A solution is given on page 58.

5.2 Coinciding birthdays

In the final problem from Subsection 1.3, you were asked to guess the probability that at least two children in a class of 24 (no twins) will have the same birthday.

This is one of the 'at least' problems that is most easily solved 'backwards', that is, using result (3.2):

$$P(E) = 1 - P(\text{not-}E).$$

If E is the event that at least two children share a birthday, then 'not-E' is the event that all the children have different birthdays. It is much easier to calculate the probability of this event directly than to calculate the probability of the actual event required.

Activity 5.8 Assumptions

What assumptions would you need to make in order to tackle this problem?

Comment

The basic assumption you need to make is that a child's birthday is equally likely to fall on any day of the year. And to keep the problem manageable, ignore leap years.

Having made the above assumption, we are ready to tackle the problem. For clarity of presentation, let us suppose that the children are listed in some way (in alphabetical order, or by age, or whatever). The first child on the list can have any day for his or her birthday – it does not matter which day.

The second child must not share a birthday with the first. Counting outcomes, 364 days out of 365 give a different birthday, so the probability that the second child does not share a birthday with the first is $\frac{364}{365}$.

Activity 5.9 The third and fourth children's birthdays

(a) If the first two children's birthdays are on different days, what is the probability that the third child does not share a birthday with either of the first two?

(b) If the first three children's birthdays are on different days, what is the probability that the fourth child does not share a birthday with any of the first three?

(c) What is the probability that none of the first four children share a birthday?

A solution for parts (a) and (b) is given on page 58. For part (c), see overleaf.

If we pick any four people at random, then on 364 occasions out of 365, the first two will have different birthdays, on $\frac{363}{365}$ of these occasions the third person's birthday will be different from the first two, and on $\frac{362}{365}$ of these occasions the fourth person will have a different birthday from any of the first three. So the proportion of times in the long run that all four will have different birthdays is

$$\frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \simeq 0.984.$$

This is the probability that all the first four children have different birthdays. Can you see a pattern forming?

Activity 5.10 The other children's birthdays

(a) What is the probability that the first five children all have different birthdays?

(b) If the first 23 children have different birthdays, what is the probability that the last child's birthday is different from the other twenty-three?

(c) Write down an expression for the probability that the birthdays of the 24 children are all different, and hence calculate the probability.

(d) What is the probability that at least two of the children share a birthday?

Comment

(a) The fifth birthday will be different from the first four on $\frac{361}{365}$ of occasions, on average, so the probability that all five birthdays are different is

$$\frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \frac{361}{365} \simeq 0.973.$$

(b) If the first 23 birthdays are all different, then there are 342 possible different days for the twenty-fourth birthday. So the probability that the twenty-fourth birthday is different from the first 23 is $\frac{342}{365}$.

(c) Hence the probability that all 24 birthdays are different is

$$\frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \frac{361}{365} \times \cdots \times \frac{342}{365} \simeq 0.462.$$

the second is different from the first

the third is different from the first two

• • •

the 24th is different from the first 23

(d) So the probability that none of the children share a birthday is less than $\frac{1}{2}$, and the probability that at least two of the children share a birthday is

$$1 - 0.462 = 0.538.$$

Was your 'guess' anywhere near this? Or did you think the chances of a shared birthday were much lower? Do you find this result surprising? You may not find it so surprising if two people in your family (an aunt, a cousin, ...) share a birthday, or if two of your friends do. And the chances are that they do!

Summary of Section 5

In this section, the ideas and techniques discussed in the earlier sections have been used to tackle the questions posed in Subsection 1.3 which were still outstanding. We hope that working through this section has helped you to consolidate your understanding of the work in this chapter.

Summary of Chapter D1

In this chapter, you have been introduced to some basic ideas of probability theory. Part of your work involved examining your own intuitions about chance events. You were asked to use the simulations part of the statistics software to investigate a number of problems and to provide data on which to base conjectures. Probability theory was then used to tackle these and other problems.

Learning outcomes

You have been working towards the following learning outcomes.

Terms to know and use

Outcome, equally-likely outcomes, event, independent events, random variable, probability function, probability distribution, frequency diagram, mean of a probability distribution, geometric distribution.

Symbols and notation to know and use

The notation $P(\,)$ for a probability.
The symbol μ for the mean of a probability distribution.
The notation $P(X = j)$ for a probability involving a random variable X.

Mathematical skills

- ◇ Identify the equally-likely outcomes of an experiment, using a systematic approach where possible.
- ◇ Use the idea of counting equally-likely outcomes to calculate the probability of an event, where appropriate.
- ◇ Use the rule $P(E) = 1 - P(\text{not-}E)$ to calculate $P(E)$ in situations where it is easier to calculate $P(\text{not-}E)$ than to calculate $P(E)$ directly.
- ◇ Use the multiplication rule for independent events to calculate probabilities.
- ◇ Calculate probabilities associated with a geometric distribution.
- ◇ Calculate the mean of a geometric distribution, given the value of the parameter p of the distribution.
- ◇ Apply the formula for the mean of a geometric distribution to solve problems concerning collecting a complete set of objects.
- ◇ Calculate probabilities in situations similar to that described in the 'Coinciding birthdays' problem.
- ◇ Look for a pattern in the results of calculations for special cases, and conjecture a general result.
- ◇ Look for a pattern in the results of simulations of special cases, and conjecture a general result.

Modelling skills

- ◇ Recognise situations where a geometric distribution may be used as a model for the probability distribution of a random variable.

Features of the statistics software to use

- ◇ Run probability simulations to investigate the behaviour of a model for a range of situations involving uncertainty.

Ideas to be aware of

- ◇ The uncertainty in a random variable can be represented by a probability distribution.
- ◇ The mean of a probability distribution is interpreted as the mean predicted by the model.

Appendix: The mean of a geometric distribution

In Section 4, you saw that the mean of a probability distribution is defined to be

This is formula (4.4).

$$\mu = \sum_j j \times P(X = j).$$

The probability function of a geometric distribution is given by formula (4.2) as

$$P(X = j) = (1-p)^{j-1}p, \quad j = 1, 2, 3, \ldots.$$

So the mean of a geometric distribution is given by

$$\mu = \sum_{j=1}^{\infty} j(1-p)^{j-1}p,$$

that is,

$$\mu = p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots. \tag{A.1}$$

If we multiply both sides by $1 - p$, we obtain

$$(1-p)\mu = (1-p)p + 2(1-p)^2p + 3(1-p)^3p + 4(1-p)^4p + \cdots. \tag{A.2}$$

Subtracting equation (A.2) from (A.1) gives

$$\mu - (1-p)\mu = p + (1-p)p + (1-p)^2p + (1-p)^3p + \cdots.$$

The left-hand side of this equation is equal to

$$\mu - \mu + p\mu = p\mu.$$

The right-hand side is the sum of the probabilities $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X = 4)$, $\ldots$ for a geometric distribution. And since X must take one or other of the values $1, 2, 3, 4, \ldots$, the sum of these probabilities is equal to 1. Therefore we have

$$p\mu = 1,$$

and hence

$$\mu = \frac{1}{p},$$

as required.

Solutions to Activities

Solution 1.3

(a) All the 52 cards are equally likely to be at the top of the pack, so

$$P(\text{ace of spades}) = \tfrac{1}{52}.$$

(b) There are 4 aces in a pack of 52 cards, so, on average, we would expect an ace to be at the top of the pack 4 times out of 52. That is,

$$P(\text{ace}) = \tfrac{4}{52} = \tfrac{1}{13}.$$

(c) There are 13 hearts in a pack of 52 cards, so

$$P(\text{heart}) = \tfrac{13}{52} = \tfrac{1}{4}.$$

Solution 1.4

(a) Each ticket is equally likely to win, and there are a million tickets.

(i) The probability of winning if you buy one ticket is $\frac{1}{1\,000\,000}$.

(ii) The probability of winning if you buy ten tickets is $\frac{10}{1\,000\,000} = \frac{1}{100\,000}$.

(b) Each of the 13 983 816 selections is equally likely to occur.

(i) The probability of winning a share of the jackpot with one selection is $\frac{1}{13\,983\,816}$.

(ii) The probability of winning with ten different selections is $\frac{10}{13\,983\,816}$.

(iii) The probability of winning with 100 different selections is $\frac{100}{13\,983\,816} \simeq 0.000\,007$.

Even with 100 different selections, the probability of winning is extremely small.

Solution 3.2

An argument identical to that used when the side showing is red can be used when the side showing is white. There are three white sides; two of these have a white side on the other side of the card, so the probability that the other side is white is $\frac{2}{3}$. It is not a fair bet: the entertainer will win in the long run on this bet too.

Solution 3.4

(a) There are eight possible equally-likely outcomes:

$$hhh,\ hht,\ hth,\ htt,\ thh,\ tht,\ tth,\ ttt.$$

Notice the systematic way that the possible outcomes of three tosses of a coin have been written down: each of the first four outcomes is h followed by one of the four possible outcomes of two tosses; and each of the second four outcomes is t followed by one of the four possible outcomes of two tosses.

(b) (i) The probability of no heads is

$$P(\text{no heads}) = \tfrac{1}{8}.$$

(ii) Seven of the eight possible outcomes include at least one head, so

$$P(\text{at least one head}) = \tfrac{7}{8}.$$

(iii) Four of the eight possible outcomes include at least two heads, so

$$P(\text{at least two heads}) = \tfrac{4}{8} = \tfrac{1}{2}.$$

(iv) Three of the outcomes contain exactly two heads, so

$$P(\text{two heads}) = \tfrac{3}{8}.$$

Solution 3.5

(a) (i) The four possible family patterns are

$$GG,\ GB,\ BG,\ BB.$$

(ii) Two of these patterns contain one girl and one boy, so the probability of one girl and one boy in a family of two children is $\frac{2}{4} = \frac{1}{2}$.

(b) (i) There are sixteen possible family patterns for families of size four:

$$\begin{array}{llll} GGGG, & GGGB, & GGBG, & GGBB, \\ GBGG, & GBGB, & GBBG, & GBBB, \\ BGGG, & BGGB, & BGBG, & BGBB, \\ BBGG, & BBGB, & BBBG, & BBBB. \end{array}$$

(ii) Six of these patterns contain two girls and two boys, so the probability that the Johnsons will achieve their ideal family is $\frac{6}{16} = \frac{3}{8}$.

(c) The Watsons are right about their chances (although their reasoning is dubious). However, the Johnsons are not right: the probability of obtaining their ideal family is only $\frac{3}{8}$, not $\frac{1}{2}$ as they suppose.

Solution 3.7

(a) Although there are only 6 partitions of 9, they are not all equally likely. Some of them can occur in more than one way; for example, 522 is not the same outcome as 252 or 225. It is important to distinguish between the three dice, for instance by imagining that they are different colours, and by always recording the score on one particular die first.

(b) The possible equally-likely outcomes which give a total of 9 are as follows.

Partitions	Outcomes					
(621)	621	612	261	216	162	126
(531)	531	513	351	315	153	135
(522)	522	252	225			
(441)	441	414	144			
(432)	432	423	342	324	243	234
(333)	333					

There are 25 possible outcomes giving a total of 9, so the probability of obtaining a total score of 9 with three dice is $\frac{25}{216} \simeq 0.116$.

(c) There are 27 possible equally-likely outcomes giving a total of 10. These are listed below.

Partitions	Outcomes					
(631)	631	613	361	316	163	136
(622)	622	262	226			
(541)	541	514	451	415	154	145
(532)	532	523	352	325	253	235
(442)	442	424	244			
(433)	433	343	334			

So the probability of obtaining a total score of 10 with three dice is $\frac{27}{216} = 0.125$.

Solution 3.8

(a) The probability that each child is a girl is $\frac{1}{2}$, independently of whether the other children are girls, so the probability that the first three children are all girls is

$$P(GGG) = \tfrac{1}{2} \times \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{8} = 0.125.$$

(b)
$$\begin{aligned} P(\text{at least one boy}) &= 1 - P(\text{no boys}) \\ &= 1 - P(\text{all girls}) \\ &= 1 - \tfrac{1}{8} \\ &= \tfrac{7}{8} = 0.875 \end{aligned}$$

(c) (i) The probability that fifteen children are all boys is

$$\left(\tfrac{1}{2}\right)^{15} = \frac{1}{32\,768}.$$

(ii) Similarly, the probability that fifteen children are all girls is

$$\left(\tfrac{1}{2}\right)^{15} = \frac{1}{32\,768}.$$

So the probability that fifteen children are all the same sex, that is, either all boys or all girls, is

$$\frac{1}{32\,768} + \frac{1}{32\,768} = \frac{2}{32\,768} = \frac{1}{16\,384}.$$

Solution 3.9

(a) If 17 out of 35 faces correspond to a girl, then, assuming that all 35 faces are equally likely to come up, the probability of a girl is $\frac{17}{35}$.

(b) The probability that three children are all girls is

$$P(\text{all girls}) = \left(\tfrac{17}{35}\right)^3 \simeq 0.115.$$

(c)
$$\begin{aligned} P(\text{at least one boy}) &= 1 - P(\text{no boys}) \\ &\simeq 1 - 0.115 \\ &= 0.885 \end{aligned}$$

Solution 3.10

(a) The total score is 3 if the score on each die is 1. The probability that the score on all three dice is 1 is

$$\tfrac{1}{6} \times \tfrac{1}{6} \times \tfrac{1}{6} = \tfrac{1}{216}.$$

(b) (i) $P(\text{total score of 3 twice}) = \frac{1}{216} \times \frac{1}{216}$
$= \frac{1}{46\,656} \simeq 2.14 \times 10^{-5}$

(ii) $P(\text{total score of 3 three times}) = \left(\frac{1}{216}\right)^3$
$\simeq 9.92 \times 10^{-8}$

(iii) $P(\text{total score of 3 four times}) = \left(\frac{1}{216}\right)^4$
$\simeq 4.59 \times 10^{-10}$

These probabilities are very small indeed. Cardano was justified in regarding the occurrence of a total score of 3 three or four times in a row as worthy of suspicion.

Solution 3.11

(a) The probability that the score on a single roll of a die is not 6 is $\frac{5}{6}$, and the scores on two rolls of a die are independent, so, by the multiplication rule, the probability that neither score is a 6 in two rolls is

$$\tfrac{5}{6} \times \tfrac{5}{6} = \tfrac{25}{36}.$$

(b) The probability of no sixes in three rolls of a die is

$$\tfrac{5}{6} \times \tfrac{5}{6} \times \tfrac{5}{6} = \tfrac{125}{216}.$$

(c) The probability of a double-six when two dice are rolled is the probability that both dice come up 6, that is,

$$\tfrac{1}{6} \times \tfrac{1}{6} = \tfrac{1}{36}.$$

(d) The probability of two double-sixes in two rolls of a pair of dice is

$$\tfrac{1}{36} \times \tfrac{1}{36} = \tfrac{1}{1296}.$$

Solution 3.12

(a) From Activity 3.11(c), the probability of obtaining a double-six is $\frac{1}{36}$ so, using (3.2), the probability of failing to get a double-six is

$$1 - \tfrac{1}{36} = \tfrac{35}{36}.$$

(b) The probability of failing to get any double-sixes in 24 rolls of a pair of dice is, using the multiplication rule (3.3),

$$\left(\tfrac{35}{36}\right)^{24} \simeq 0.509.$$

(c) Hence, using (3.2), the probability of getting at least one double-six in 24 rolls of a pair of dice is

$$1 - P(\text{no double-sixes}) = 1 - \left(\tfrac{35}{36}\right)^{24} \simeq 0.491.$$

The Chevalier was correct to suspect that the second wager was not a good one. (He must have gambled a lot to detect such a small difference in probabilities!)

Solution 4.3

(a) It is clear from Figure 4.2 that the most likely number of rolls needed to obtain a six is 1; and we have already calculated its probability in Activity 4.2:

$$P(X = 1) = \tfrac{1}{6}.$$

(b) The probability that exactly five rolls are needed is, using formula (4.1) with $j = 5$,

$$P(X = 5) = \left(\tfrac{5}{6}\right)^4 \times \tfrac{1}{6} \simeq 0.080.$$

The probability that exactly ten rolls are needed is

$$P(X = 10) = \left(\tfrac{5}{6}\right)^9 \times \tfrac{1}{6} \simeq 0.032.$$

You are more likely to obtain your first six on your fifth roll than on your tenth roll.

Solution 4.4

(a) (i) The probability that the first trial results in a success is p, so

$$P(X = 1) = p.$$

(ii) The probability that the first trial is a failure is $1 - p$.

(b) $P(X = 2) = P(\text{failure}) \times P(\text{success}) = (1 - p)p$

(c) $P(X = 3) = P(\text{failure}) \times P(\text{failure}) \times P(\text{success})$
$= (1 - p)^2 p$

(d) Continuing the pattern, we obtain the following expression for $P(X = j)$:

$$\underbrace{P(\text{failure}) \times \cdots \times P(\text{failure})}_{j-1 \text{ failures}} \times \underbrace{P(\text{success})}_{\text{followed by a success}}$$
$$= (1 - p)^{j-1} p.$$

Solution 4.5

(a) The size of such a family has a geometric distribution, so the most likely size is 1. (In this case, $p = P(\text{success}) = \frac{1}{2}$.)

(b) If X is the number of children they have, then the probability that they will have only one child is

$$P(X = 1) = \tfrac{1}{2}.$$

(c) The probability that they will have three children is

$$P(X = 3) = (1 - p)^2 p = \left(\tfrac{1}{2}\right)^2 \times \tfrac{1}{2} = \tfrac{1}{8}.$$

(d) The probability that they will have six children is

$$P(X = 6) = (1 - p)^5 p = \left(\tfrac{1}{2}\right)^5 \times \tfrac{1}{2} = \tfrac{1}{64}.$$

Solution 4.6

(a) The probability that more than ten rolls are needed to obtain a six is just the probability that none of the first ten rolls is a six:

$$P(X > 10) = \left(\tfrac{5}{6}\right)^{10} \simeq 0.162.$$

(b) The probability that at most ten rolls are needed is, using result (3.2),

$$\begin{aligned} P(X \le 10) &= 1 - P(X > 10) \\ &\simeq 1 - 0.162 = 0.838. \end{aligned}$$

(c) The probability that more than twenty rolls of the die are needed is

$$P(X > 20) = \left(\tfrac{5}{6}\right)^{20} \simeq 0.026.$$

Solution 4.7

(a) The probability that the Smiths will have more than four children is just the probability that the first four children are all boys, that is,

$$P(X > 4) = \left(\tfrac{1}{2}\right)^4 = \tfrac{1}{16}.$$

(b) The probability that they will have four children or fewer is, using result (3.2),

$$P(X \leq 4) = 1 - P(X > 4) = \tfrac{15}{16}.$$

Solution 4.8

(a) Since $p = \frac{1}{6}$, the average (that is, the mean) number of rolls of a die needed to obtain a six is $1/\frac{1}{6} = 6$.

(b) Since $p = \frac{1}{2}$, the average size of family for couples like the Smiths is $1/\frac{1}{2} = 2$.

(c) Since $p = \frac{2}{9}$, the average number of darts needed to hit the bull's-eye is $1/\frac{2}{9} = \frac{9}{2} = 4\frac{1}{2}$.

Note that a mean does not have to be a whole number. This result just says that the average number of darts Tom needs to throw to hit the bull's-eye is between 4 and 5.

Solution 5.7

This problem is equivalent to collecting a complete set of size 2, the two 'objects' being a girl and a boy.

The first child is certain to be either a girl or a boy. Thereafter, the probability that the next child is of the other sex is $\frac{1}{2}$, so the mean number of additional children needed to obtain a child of the other sex is $1/\frac{1}{2} = 2$.

Hence the mean total number of children needed to obtain at least one child of each sex is $1 + 2 = 3$. The mean size of 'completed' families is 3.

A course team member carried out 100 simulations. The sizes of the 'completed' families varied between 2 and 8, and the average family size was 3.03, quite close to the theoretical mean family size.

Solution 5.9

(a) There are 363 possible different days for the third child's birthday, so the probability that it is on a different day from the first two children's birthdays is $\frac{363}{365}$.

(b) There are 362 possible different days for the fourth child's birthday, so the probability that it is on a different day from the first three children's birthdays is $\frac{362}{365}$.

Solutions to Exercises

Solution 3.1

(a) There are two tickets ending in either 0 or 5 in each group of ten consecutively numbered tickets. Since there are 50 groups of ten tickets in the 500 tickets, there are 50×2 winning tickets; that is, there are 100 tickets with a number ending in 0 or 5.

(b) Using formula (3.1),

$$P(\text{win}) = \tfrac{100}{500} = \tfrac{1}{5}.$$

Solution 3.2

(a) There are 4 possible outcomes for the score on the first die and 4 possible outcomes for the score on the second die. So there are $4 \times 4 = 16$ possible outcomes when the two dice are rolled together. A total of 5 can be obtained in four ways:

1 on the first die, 4 on the second die;
2 on the first die, 3 on the second die;
3 on the first die, 2 on the second die;
4 on the first die, 1 on the second die.

So, using formula (3.1),

$$P(\text{total of } 5) = \tfrac{4}{16} = \tfrac{1}{4}.$$

(b) The probability that the first die lands on a 2 is $\frac{1}{4}$, and the probability that the second die lands on an odd number is $\frac{2}{4} = \frac{1}{2}$. The scores on the two dice are independent, so, using the multiplication rule (3.3), the probability that the first die lands on a 2 and the second die lands on an odd number is

$$\tfrac{1}{4} \times \tfrac{1}{2} = \tfrac{1}{8}.$$

Solution 3.3

(a) The probability that the score on a single roll of a tetrahedral die is 4 is $\frac{1}{4}$, so, by the multiplication rule (3.3), the probability that both dice land on 4 is $\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$.

(b) Using (3.2), the probability of failing to obtain a double-four when the two dice are rolled is

$$1 - P(\text{double-four}) = 1 - \tfrac{1}{16} = \tfrac{15}{16}.$$

(c) The probability of failing to obtain any double-fours in six rolls of a pair of tetrahedral dice is, using the multiplication rule (3.3),

$$\left(\tfrac{15}{16}\right)^6 \simeq 0.679.$$

(d) Hence, using (3.2), the probability of obtaining at least one double-four in six rolls of a pair of tetrahedral dice is

$$\begin{aligned} 1 - P(\text{no double-fours}) &= 1 - \left(\tfrac{15}{16}\right)^6 \\ &\simeq 1 - 0.679 \\ &= 0.321. \end{aligned}$$

Solution 4.1

The probability that each child born is a girl is $\frac{17}{35}$, so $p = P(\text{success}) = \frac{17}{35}$.

(a) If X is the number of children that a couple such as the Smiths have, then the probability that they have four children is

$$P(X = 4) = (1-p)^3 p = \left(\tfrac{18}{35}\right)^3 \times \tfrac{17}{35} \simeq 0.066.$$

(b) The probability that they have more than four children, $P(X > 4)$, is given by

$$P(\text{first four children are boys}) = \left(\tfrac{18}{35}\right)^4 \simeq 0.070.$$

(c) The average family size for such couples is

$$\frac{1}{p} = \frac{1}{17/35} = \tfrac{35}{17} \simeq 2.06.$$

That is, the average number of children in such families is just over 2. (Recall that the mean does not have to be a whole number.)

Solution 4.2

(a) The probability of a 'success' in a single week is $p = \frac{1}{13\,983\,816}$, so the average time people have to wait to win a share of the jackpot is

$$\begin{aligned} \frac{1}{p} = 13\,983\,816 \text{ weeks} &\simeq 13\,983\,816/52 \text{ years} \\ &\simeq 269\,000 \text{ years}, \end{aligned}$$

that is, over a quarter of a million years!

(b) The number of weeks in 50 years is approximately $50 \times 52 = 2600$, so we require the probability of failing to win a share of the jackpot for 2600 weeks in a row. This is

$$\left(\frac{13\,983\,815}{13\,983\,816}\right)^{2600} \simeq 0.9998,$$

so you almost certainly will not win, even if you make a selection every week for 50 years!

Index

The Open University
ISBN 0 7492 6661 9